

Washington University in St. Louis
Washington University Open Scholarship

All Theses and Dissertations (ETDs)

12-26-2013

Systems Metabolic Engineering of Microbial Cell Factories for the Synthesis of Value-added Chemicals

Arul Mozhy Varman
Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

 Part of the [Engineering Commons](#)

Recommended Citation

Varman, Arul Mozhy, "Systems Metabolic Engineering of Microbial Cell Factories for the Synthesis of Value-added Chemicals" (2013). *All Theses and Dissertations (ETDs)*. 1176.
<https://openscholarship.wustl.edu/etd/1176>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering and Applied Science
Department of Energy, Environmental, and Chemical Engineering

Dissertation Examination Committee:

Yinjie J. Tang, Chair
Himadri B. Pakrasi, Co-chair
Pratim Biswas
Gautam Dantas
Tae Seok Moon
Venkat Subramanian
James Umen

Systems Metabolic Engineering of Microbial Cell Factories for the
Synthesis of Value-added Chemicals

by

Arul Mozhy Varman

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2013

Saint Louis, Missouri

© 2013, Arul Mozhy Varman

Table of Contents

List of Figures	v
List of Tables	ix
Acknowledgements	x
Abstract	xiii
Chapter 1: Introduction to systems metabolic engineering of microbial cell factories	1
1.1 Introduction.....	1
1.2 Microbial cell factories	2
1.3 Tools for genetic engineering of microbial hosts.....	5
1.4 Cyanobacteria as a microbial cell factory	6
1.5 Modeling and Systems Analysis	10
Chapter 2: Statistics-based model for prediction of chemical biosynthesis yield from <i>Saccharomyces cerevisiae</i>	13
2.1 Background.....	14
2.2 Model development	16
2.3 Results and discussion	19
2.4 Model applications and limitations	23
2.5 Comparison to the previously published <i>E. coli</i> model	25
2.6 Conclusions.....	26
Acknowledgments.....	27
Chapter 3: The use of ¹³C-based analysis to elucidate the intrinsic biosynthesis yields	40
3.1 Introduction.....	41
3.2 Product yield using rich medium	42
3.3 Product yield during co-metabolism of multiple carbon substrates.....	44
3.4 Accurate laboratory analysis of product concentrations	45
3.5 Assessment of maximum product yield	47
3.6 Product yield from unconventional engineered pathway.....	50
3.7 Conclusion	53
Acknowledgement	54
Chapter 4: Metabolic engineering of <i>Synechocystis</i> 6803 for isobutanol production	55
4.1 Introduction.....	56

4.2 Materials and methods	57
4.2.1 Chemicals and reagents.....	57
4.2.2 Culture medium and growth conditions.....	57
4.2.3 Plasmid construction and transformation of <i>Synechocystis</i> 6803.....	58
4.2.4 Reverse transcription PCR (RT-PCR).....	59
4.2.5 Isobutanol quantification assay.....	60
4.2.6 ¹³ C-experiment to detect carbon contribution of glucose.....	61
4.3 Results.....	62
4.3.1 Construction of an isobutanol producing <i>Synechocystis</i> 6803 strain.....	62
4.3.2 Isobutanol synthesis under autotrophic and mixotrophic growth.....	63
4.3.3 In situ alcohol concentrating system using a solvent trap.....	64
4.4 Discussion.....	65
Acknowledgment.....	67
Chapter 5: Photoautotrophic production of D-lactic acid in an engineered cyanobacterium.....	75
5.1 Background.....	77
5.2 Results and Discussion.....	79
5.3 Conclusions.....	83
5.4 Materials and methods.....	84
5.4.1 Chemicals and reagents.....	84
5.4.2 Medium and growth conditions.....	84
5.4.3 Plasmid construction and transformation.....	85
5.4.4 D (-) lactate analysis.....	86
5.4.5 ¹³ C isotopomer experiment.....	86
Acknowledgements.....	87
Chapter 6: Kinetic modeling and isotopic investigation of isobutanol fermentation by two engineered <i>Escherichia coli</i> strains.....	98
6.1 Introduction.....	99
6.2 Materials and Methods.....	100
6.2.1 Pathway construction.....	100
6.2.2 Fermentation conditions.....	101
6.2.3 Analytical methods for biomass and metabolites.....	102
6.2.4 ¹³ C-experiments for analyzing nutrient contributions to isobutanol productions.....	102
6.2.5 Model formulation.....	103

6.3 Results and discussion	106
6.3.1 Isobutanol fermentation results	106
6.3.2 Kinetic modeling of isobutanol fermentation.....	107
6.3.3 Analysis of the role of yeast extract for isobutanol synthesis	109
6.4 Concluding remarks	110
Acknowledgement	110
Chapter 7: Conclusions and perspectives.	121
7.1 Summary	121
7.2 Challenges with commercialization of industrial biotechnology.....	123
7.3 Challenges with cyanobacterial bioprocessing.	125
7.4 Recent developments in synthetic biology.....	126
7.5 Future directions	128
References.....	130
Appendix 1: Microbial metabolisms and cell culture models for biofuel production. Bioenergy: Principles and Applications.	144
Appendix 2: Engineering <i>Escherichia coli</i> to convert acetic acid to free fatty acids.	165
Appendix 3: Evaluating factors that influence microbial synthesis yields by regression with numerical and categorical variables.....	176
Curriculum vitae.....	186

List of Figures

- Figure 1.1:** Metabolic engineering pathways for biofuel production.....3
- Figure 1.2:** Schematic representation of engineered biochemical pathways for the production of biofuels in cyanobacteria. Abbreviations: RuBP, Ribulose-1,5-bisphosphate3-PGA; 3-phosphoglycerate; Kdc, ketoacid decarboxylase; ADH, alcohol dehydrogenase; PDC, pyruvate decarboxylase; ACC, acetyl-CoA carboxylase; AAR, acyl-ACP reductase; AAD, aldehyde decarboxylase; FAR, fatty acyl-CoA reductase; Ter, trans-2-enoyl-CoA reductase; Hbd, 3-hydroxybutyryl-CoA dehydrogenase; AdhE2, aldehyde/alcoholdehydrogenase. Highlighted reactions indicate the pathways focused in this study... ..9
- Figure 1.3:** Metabolic network modeling and analysis. FBA profiles the "optimal" metabolism for the desired performance (can be genome-scale); ¹³C-MFA measures *in vivo* operation of the central metabolic network (< 100 reactions).12
- Figure 2.1:** Metabolic pathway for the biosynthesis of major products. The blue box represents central metabolism and the yellow box represents secondary metabolism. Solid arrows signify single step reaction and dotted arrow signify multiple steps. Abbreviations: ACoA – Acetyl-CoA; DAP – Dihydroxyacetone-Phosphate; DAHP – 3-Deoxy-D-Arabino-Heptulosonate-7-Phosphate; DHA – Dihydroxyacetone; F6P - Fructose-6-Phosphate; FBP – Fructose 1,6-bisphosphate; G6P – Glucose-6-Phosphate; GADP – Glyceraldehyde-3-Phosphate; Oxa – Oxaloacetate; Oxo – 2-Oxoglutarate; PEP – Phosphoenolpyruvate; PHB – Poly[(R)-3-hydroxybutyrate]; pHCA – p-Hydroxycinnamic acid; R5P – Ribose-5-Phosphate; Ru5P – Ribulose-5-Phosphate; Suc – Succinate; X5P – Xylulose-5- Phosphate.28
- Figure 2.2:** Model results. A) Plot of the actual logarithmic yields against the logarithmic yields generated by the regression model. The line drawn as diagonal to the plot is one-to-one and passes through the origin. The data points have an R² value of 0.55. B) Plot of residuals against fitted values. C) Model validation using newly published data (2010~2011) 1 - β-amyrin; 2 - ascorbic acid; 3 – monoterpene; 4 – vanillin; 5 - succinic acid31
- Figure 2.3:** *S. cerevisiae* model prediction of biosynthesis yields for other industrial yeast species..... 32
- Figure 3.1:** Schematic description of microbial metabolism. Many microbes have the ability to co-metabolize diverse feedstock. The enrichment of labeling in the product acts as an indicator for the relative uptake flux of sugars42
- Figure 3.2:** (A) Biosynthesis yield analyzed by feeding cells with ¹³C-substrates (B) Relative product yields from a primary substrate (a – Isobutanol from glucose in a low performance strain; ab – valine from glucose in a low performance strain; b – Isobutanol from glucose in JCL260; bb – valine from glucose in JCL260) ; c – Free fatty acids from acetate in an *E.coli* strain ; d - biomass from glucose in wild type *Synechocystis* 6803 ; e - D-lactate from acetate in engineered *Synechocystis* 6803 . Relative yield is calculated based on ¹³C concentrations in

- the final product. Abbreviations: GAP, Glyceraldehyde -3- phosphate; PYR, pyruvate; KIV, ketoisovalerate.....44
- Figure 3.3:** Schematic showing the dynamics of product concentration. C_a is the actual concentration and C_m is the measured product concentration. C^* indicates the concentration of the ^{13}C product added as an internal standard.....47
- Figure 3.4:** (A) Theoretical Yield as a function of P/O ratio and ATP maintenance without biomass growth. (B) Theoretical Yield as a function of P/O ratio and ATP maintenance at growth rate $v(8)=3.6$. The units of yield and ATP maintenance are ‘g C16:0 fatty acid/g glucose’ and ‘mol ATP /g glucose’ respectively. The infeasible range in the surface plot indicates that, energy cannot be balanced for fatty acid or biomass production in that region, resulting in zero yield50
- Figure 3.5:** ^{13}C analysis to study the carbon assimilation during syngas fermentation ($^{13}\text{CO}_2$, ^{12}CO and H_2). Analysis of metabolite labeling patterns can determine CO_2 and CO utilization for pyruvate production. The isotopomer data of pyruvate were used as a demonstration of ^{13}C applications for product yield calculations.....51
- Figure 3.6:** Threonine and citramalate pathway for the synthesis of 1-butanol. The carbon rearrangement network shows the labeling of 1-butanol for both the pathways, when fed with 1- ^{13}C pyruvate and ^{13}C bicarbonate.....52
- Figure 4.1:** (a) Schematic representation to show the integration of the genes *kivd* and *adhA* into the genome of *Synechocystis* 6803. Colony PCR performed to verify the integration of the insert into the genomic DNA of the mutant (AV03). The vector *ptka3* was used as a template for the positive control and wild-type cells were used as negative control. Colony PCR of AV03, showed the presence of a band (8.3kb) the same size as the positive control (+ve) and the absence of the negative control (WT) band. (b) IB synthesized by engineered *Synechocystis* 6803 under different IPTG concentrations (n=3). (Inset) Result of an RT-PCR performed to detect the expression of the heterologous genes *kivd* (Top: 500bp from *kivd*) and *adhA* (Bottom: 200bp from *adhA*). Lane 1, wild-type 6803 (WT); Lane 2, AV03 with 0 mM IPTG; Lane 3, AV03 with 0.5mM IPTG; Lane 4, AV03 with 1mM IPTG.68
- Figure 4.2:** (a) Growth curves of *Synechocystis* 6803 WT and AV03 (n=3, shake flask cultures): \diamond WT under autotrophic, \blacklozenge WT under mixotrophic, \circ AV03 under autotrophic and \bullet AV03 under mixotrophic conditions (note: growth curve of AV03 under mixotrophic condition overlaps with autotrophic growth curves of AV03 and WT). (b) IB synthesized in AV03 under autotrophic conditions (only HCO_3), heterotrophic (only glucose) and mixotrophic (both HCO_3 and glucose) conditions (n=3, shake flask cultures with closed caps). (c) Percentage carbon contribution of glucose for synthesizing amino acids and isobutanol in the wild-type (WT) and the mutant strain (AV03) as measured on day 9 (shake flask cultures with closed caps). Isotopomer analysis (TBDMS based method) of proteinogenic amino acids confirms the low ^{13}C -glucose utilization by the mutant. The error bar represents the 2% technical error of the instrument.70

- Figure 4.3:** Toxic effects of IB on the growth of *Synechocystis* 6803. IB was added to a final concentration (g/L, n=2) of \diamond 0, \square 0.2, Δ 0.5, \circ 1, \blacksquare 2 and \bullet 5 to a *Synechocystis* 6803 culture with an initial OD₇₃₀ ~ 0.8.71
- Figure 4.4:** Net concentration of IB synthesized (columns) and biomass growth (curves) by the AV03 culture under different conditions (n=3). a – IB with 0.5 mL oleyl alcohol (Autotrophic); b – IB with 1 mL oleyl alcohol (Autotrophic); c – IB with 0.5 mL oleyl alcohol and glucose (Mixotrophic); d – IB with no oleyl alcohol (Autotrophic, negative control); a1 – OD₇₃₀ with 0.5 mL oleyl alcohol (Autotrophic); a2 – OD₇₃₀ with 1 mL oleyl alcohol (Autotrophic); a3 – OD₇₃₀ with 0.5 mL oleyl alcohol and glucose (Mixotrophic); a4 – OD₇₃₀ with no oleyl alcohol (Autotrophic, negative control). (Inset) Schematic representation of the in situ IB removal system used to increase the production of IB.72
- Figure 5.1:** Metabolic engineering of *Synechocystis* 6803 for the synthesis of D-lactic acid. (A) Metabolic pathway for D-lactate synthesis. Lactate permeation through the cell membrane occurs either via a lactate transporter or by passive diffusion; Red arrows indicate the heterologous pathway engineered into *Synechocystis* 6803. Abbreviations: GlyDH*, mutant glycerol dehydrogenase; TH, Transhydrogenase; 3PGA, 3-phosphoglycerate; CoA, Coenzyme A; G1P, glucose 1-phosphate; F6P, fructose 6-phosphate; PHB, poly- β -hydroxybutyrate; RuBP, ribulose 1,5-bisphosphate. (B) Colony PCR to verify the presence of the heterologous genes of the mutant glycerol dehydrogenase (Left picture) and transhydrogenase (Right picture) in the engineered strains of *Synechocystis* 6803. *gldA101* was amplified with primers *gldA-o-F3* and *gldA-o-R*; *gldA101-syn* was amplified with primers *gldA-o-F* and *gldA-o-R2*; *sth* was amplified with primers *tranNADH-F* and *tranNADH-R*91
- Figure 5.2:** Autotrophic production of D-lactate in the engineered strains of *Synechocystis* 6803. (A) Growth and (B) D-lactate production in the engineered strains (n = 3). Circles: AV08 (with *gldA101*). Triangles: AV10 (with *gldA101-syn* and *sth*) and Squares: AV11 (with *gldA101-syn*)92
- Figure 5.3:** Mixotrophic production of D-lactate by AV10. (A) Growth and (B) lactate production in the engineered *Synechocystis* 6803 strain AV10 (n = 3), with the provision of additional organic carbon source, i.e., with glucose and acetate (Mixotrophic metabolism). Squares: with acetate. Circles: with glucose.....93
- Figure 5.4:** Isotopomer analysis showing the mass fraction of isotopomers for selected proteinogenic amino acids [TBDMS based] and D-lactate [MSTFA based]. Standard abbreviations are used for amino acids in the figure. (A) Cultures grown with 5 g/L of [1, 2-¹³C] glucose and (B) Cultures grown with 15 mM of [1,2-¹³C] acetate. “white bar” m0 – mass fraction without any labeled carbon; “grey bar” m1 – mass fraction with one labeled carbon; “black bar” m2 – mass fraction with two labeled carbon. (Note: natural ¹³C makes up about 1.1% of total carbon as measurement background).....95

Figure 5.5: Nucleotide sequence alignment of <i>gldA101</i> and the codon-optimized <i>gldA101</i> (i.e., <i>gldA101-syn</i> , synthesized by Genewiz Inc). Conserved nucleotide sequences in <i>gldA101-syn</i> are indicated as dotted lines	96
Figure 5.6: Autotrophic growth curve for <i>Synechocystis</i> 6803 strains shows similar growth of the engineered D-lactate producing strains as compared to the wild type strain. Diamond: Wild type. Square: AV08. Triangle: AV10. Circle: AV11.....	97
Figure 5.7: Mass spectra obtained via GC-MS confirm the presence of lactate in the cell culture supernatant of AV10 strain. D/L lactate enzyme kit (R-Biopharm) was used to further confirm that the product is an optically pure D-lactate.....	97
Figure 6.1: Metabolism in the <i>E. coli</i> strains for IB production. R_X , $R_{X,YE}$, R_A , R_E , R_L , and R_{IB} were shown in the Equations 6.1~6.12. IB synthesis consumes one mole NADPH (by keto-acid reductoisomerase) and one mole NADH (by aldehyde reductase). The cell metabolism removes the redundant NADH by O_2 oxidization or by synthesis of lactate and ethanol.....	113
Figure 6.2: Growth kinetics after IPTG induction (F1). The circles were experimental measurements, and the solid lines were simulations from the Monod kinetic model	114
Figure 6.3: Growth kinetics after IPTG induction (F2).....	115
Figure 6.4: Growth kinetics after IPTG induction (F3, biomass growth data were from two identical batch experiments).	116
Figure 6.5: Growth kinetics after IPTG induction (F4).	117
Figure 6.6: Growth kinetics after IPTG induction (F5).The dotted line was model prediction of IB concentrations without gas stripping.	118
Figure 6.7: The fraction of ^{13}C carbon in metabolites from the low performance (A) and JCL260 (B) IB-producing strains. The biomass was grown on fully labeled ^{13}C -glucose, with 1 g/L (black bar) or 5 g/L (gray bar) nonlabeled yeast extract (n=2, GC-MS standard errors < 2%).	120
Figure 7.1: Schematic diagram of the various factors that play a major role in the economic feasibility of a biofuel production process	124
Figure 7.2: Schematic diagram of an integrated iterative approach required for the development of high-performance microbial strains towards industrial commercialization.....	129

List of Tables

Table 1.1: Commonly employed microbes for biofuel production.....	4
Table 2.1: Ordinal variables used in the linear regression model	33
Table 2.2: Dataset used for linear regression.....	34
Table 2.3: Regression coefficients and P-values for <i>S. Cerevisiae</i> Model	39
Table 3.1: Simplified biochemical reactions considered in the model	49
Table 4.1: Metabolic engineering of cyanobacterial strains for biofuel production.	73
Table 4.2: Primer sequences used in this study.	74
Table 5.1: Primer sequences.....	88
Table 5.2: Plasmids and strains.....	89
Table 6.1: Recent studies on biobutanol production by engineered microorganisms.....	111
Table 6.2: Parameters of Monod model for <i>E. coli</i> IB fermentation.....	112

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Yinjie J. Tang for providing me with an excellent opportunity to be a part of his research group. I also thank him for the motivation, financial support and the close guidance he has provided me throughout my doctoral research. I sincerely feel that it was his influence that has brought the best out of me and working in his lab would be one of the most fruitful years of my life. He gave the freedom to explore my ideas as well as provided me with the much needed support to take the projects to completion. I wish I could work tirelessly like him all through the week.

I would also like to thank my co-advisor Prof. Himadri Pakrasi for providing me with a wonderful opportunity to work and learn from his lab. I also thank him for his mentoring all throughout my presence at Washington University. My special thanks to Prof. Pratim Biswas for providing me his support and his guidance from the outset. The meetings with him though sporadic, have always made me walk out of his office with a good feel and his words will always be engraved in my memory. I also thank my dissertation committee members, Dr. Gautam Dantas, Dr. Venkat Subramanian, Dr. James Umen, and Dr. Tae Seok Moon for giving their valuable suggestions, critical comments and support at different stages of my research.

I am grateful to the McKelvey family for the fellowship they provided me in my first year. I also convey my appreciation to all the administrative staff in the department of EECE - Rose, Trisha, Tim, Beth, Kim, Lesley, Lynn and Patty for making my life at EECE easier. I sincerely acknowledge the financial support provided by the funding agencies National Science Foundation (MCB0954016), U.S. Department of Energy, Consortium for Clean Coal Utilization

and the bridge fund at Washington University. I also thank Dr. K.T Shanmugam from the University of Florida for offering us the plasmid pQZ115 and Dr. James Liao from University of California Los Angeles for offering us the JCL260 strain without any delay.

I am grateful to my mentors Dr. Abhay Singh and Dr. Anindita Bandyopadhyay from the Pakrasi lab. I appreciate the time they spent in teaching me the experimental skills required to carry out metabolic engineering in cyanobacteria. I was also blessed to have had Professors who were truly inspiring and I would like to mention two of them from India who have left a lasting impression on me, Dr. S. Sundaramoorthy and Dr. Vinay A. Juvekar. I will also cherish the excellent discussions I had with Dr. Yi Xiao in Tang lab and he did impart a different perspective of metabolic engineering thinking within me. I should also mention the eventful research partnership I had with Dr. Yu Yi from Wuhan University and we had wonderful moments together apart from the exciting research results we had achieved together.

I would like to express my thanks to all the members of the Pakrasi lab and I have learnt a lot from every one of them. In a similar vein, many thanks to past and present members of the Tang lab, Dr. Bing Wu, Dr. Xueyang Feng, Lian He, You Le, Gang Wu and Whitney Hollinshead. Each one of them had offered their valuable help whenever needed and I also appreciate the contributions they made as co-authors without complaints. I appreciate the contribution of the various rotation and summer students, Yogesh Goyal, Peter Colletti, Stephanie Deng, Amelia Nguyen, Cheryl Immethun, Amelia Chen, Diany Li and Zach Hembree to my project works all throughout my presence in the Tang lab. I enjoyed teaching and passing on the little knowledge I had borrowed from the seniors before me.

Finally, I owe a special thanks to my friends and family members. Thanks to my parents and my in-laws for always believing in me and being patient enough throughout this whole journey. Their prayers, love, support and encouragement has made me travel this far. My friends have always been inspiring with the discussions we had and those discussions always strengthened the zeal in me to pursue my PhD. I would like to recognize a few of my friends from a long list of them, who offered their support and encouragement at every crucial hurdles I had faced in my life, Murali Krishna, Prateek, Parthiban Jayaraman and Raj Barath Janarthanan. I would like to thank my wife Kanimozhi, for her unconditional support, for giving a pat on my back whenever I needed and for making our home a very warm place to enter. Thanks for being with me all through the late night lab experiments and for editing and proofreading my works. Above all, I would like to thank God for the every blessing he has bestowed upon me. I would like to dedicate this thesis to all my loved ones.

Aud Moxby Varman

Abstract

Systems Metabolic Engineering of Microbial Cell Factories for the
Synthesis of Value-added Chemicals

by

Arul Mozhy Varman

Doctor of Philosophy in Energy, Environmental, and Chemical Engineering

Washington University in St. Louis, 2013

Professor Yinjie J. Tang, Chair

Microbial cell factories offer us an excellent opportunity for the conversion of many different cheaply available raw materials into valuable chemicals. Systems metabolic engineering aims at developing rational strategies for the engineering of microbial hosts by providing global level information of a cell. This dissertation focuses on metabolic engineering, bioprocess modeling and pathway analysis, to develop robust microbial cell factories for the synthesis of value-added chemicals. The following research tasks were completed in this regard.

First, statistical models were developed for the prediction of product yields in engineered microbial cell factories - *Saccharomyces cerevisiae* and *Escherichia coli* (Chapter 2). A large space of experimental data for chemical production from recent references was collected and a statistics-based model was developed to calculate production yield. The input variables (numerical or categorical variables) for the model represented the number of enzymatic steps in the biosynthetic pathway of interest, metabolic modifications, cultivation modes, nutrition and oxygen availability. In addition, the use of ^{13}C -isotopomer analysis method was proposed for the accurate determination of product yields in engineered microbes under complex cultivation conditions (Chapter 3).

Second, metabolic engineering of the cyanobacterium, *Synechocystis* sp. PCC 6803 was performed for synthesizing isobutanol under phototrophic conditions (Chapter 4). With the expression of the heterologous genes from the Ehrlich Pathway, by incorporating an *in situ* isobutanol harvesting system, and also by employing mixotrophic conditions, the engineered *Synechocystis* 6803 strain accumulated a maximum of ~300 mg/L of isobutanol in a 21 day culture. In addition, *Synechocystis* 6803 was engineered for the synthesis of D-lactic acid (Chapter 5), via overexpression of a novel D-lactate dehydrogenase (encoded by *gldA101*). The production of D-lactate was further improved by employing three strategies: (i) cofactor balancing, (ii) codon optimization, and (iii) process optimization. The engineered *Synechocystis* 6803 produced 2.2 g/L D-lactate under photoautotrophic conditions with acetate, the highest reported lactate titer among all known cyanobacterial strains.

Finally, an *E. coli* cell factory was engineered to study the fermentation kinetics for scaled-up isobutanol production (Chapter 6). Through kinetic modeling (to describe the dynamics of biomass, products and glucose concentration) and isotopomer analysis, we have also offered metabolic insights into the performance trade-off between two engineered isobutanol producing *E. coli* strains (a high performance and a low performance strain). The kinetic model can also predict isobutanol production under different fermentation conditions. I and my colleagues have also demonstrated that *E. coli* cell factory can also be used for converting waste acetate into free fatty acids through metabolic engineering. In conclusion, the opportunities and commercial limitations with current biotechnology as well as the role of systems metabolic engineering for the development of high performance microbial cell factories were discussed (Chapter 7).

Chapter 1: Introduction to systems metabolic engineering of microbial cell factories

1.1 Introduction

Biomass-derived carbon and energy have been used by human society for a long time. This dependence was shifted to petroleum derived carbon and energy in recent times. A 2008 census indicated that most of the energy utilized worldwide came from the burning of fossil fuels and it accounted for about 80% of the energy consumed¹. The U.S. Energy Information Administration had projected a 49% increase in global energy demand from 2007 to 2035². This dependence on fossil fuels cannot go on forever as oil reserves have started dwindling. Furthermore, the USEPA reports that the atmospheric CO₂ concentrations has increased by up to 35% since the industrial revolution in the 1700's³, while CO₂ produced from burning fossil fuels contributed to about 56.6% (2004 data) of the total greenhouse gas emissions³. In consideration of the energy security and environmental concerns there is a growing need for the production of biofuels and petroleum-derived chemicals from renewable sources.

For the production of chemicals from microbes to be economical, the target chemical must be produced at high yield, titer and productivity. These traits are difficult to be met by naturally occurring microbes⁴. Henceforth, microbes must be engineered to achieve the desired traits. With the advent of recombinant DNA technology, we now have the tools to redesign metabolic pathways for the production of chemicals from renewable materials. Technologies beyond simple genetic engineering are often required to achieve a desired phenotype and much of this rational modification has been performed in the form of metabolic engineering. Metabolic engineering is the improvement of cellular activities by manipulation of enzymatic, transport,

and regulatory functions of the cell by the application of recombinant DNA technology⁵. On the other hand, systems biology Metabolic engineering can be integrated with systems analysis and modeling to perform rational engineering of microbial hosts⁶.

1.2 Microbial cell factories

Microbial biocatalysts offer several advantages in producing small-molecule chemicals. Unlike conventional chemical syntheses which are heavily dependent on petroleum-derived substrates, microbes are able to use renewable materials to synthesize many commodity chemicals and fuels⁷. Due to its scalability, microorganisms are also suitable platforms to synthesize pharmaceutical molecules. Among the many industrial microorganisms, *Escherichia coli* and *Saccharomyces cerevisiae* have long been the industrial workhorses preferred for metabolic engineering applications. These two organisms have well-established genetic tools and have been explored to create industrial scale production of chemicals from microbes⁸. Developments in genetic tools have led to the ability to efficiently engineer *E. coli* as a biocatalyst for the production of a wide variety of chemicals, potential biofuels and pharmaceuticals⁹. *S. cerevisiae* is typically known for its robustness in fermenting sugars into alcohol. It has also gained importance as a heterologous platform to synthesize many precursors of commodity chemicals and pharmaceuticals⁷.

Sugars (such as glucose, xylose starch, and sucrose) have been widely used for biofuel production, which can be obtained either from food crops (corn, sugarcane, sugar beet) or from biomass polymers (i.e., cellulose and hemicellulose). To reduce feedstock costs, a great deal of effort has been focused on the isolation, characterization and engineering of a handful of species (e.g., *Clostridium thermocellum* and *Clostridium phytofermentans*) that can utilize cheap

biomass for bioproduct synthesis (such as ethanol)⁸. Engineered *Clostridium cellulolyticum* has been shown to produce isobutanol directly from crystalline cellulose¹⁰. More recently, *E. coli* was engineered for the production of biodiesel directly from hemicelluloses, a plant derived biomass¹¹. Utilizing non-sugar-based substrates, such as glycerol, lactate, acetate, CO₂, and syngas (CO, CO₂ and H₂), for the production of value added chemicals has been a trend in recent years (Figure 1.1).

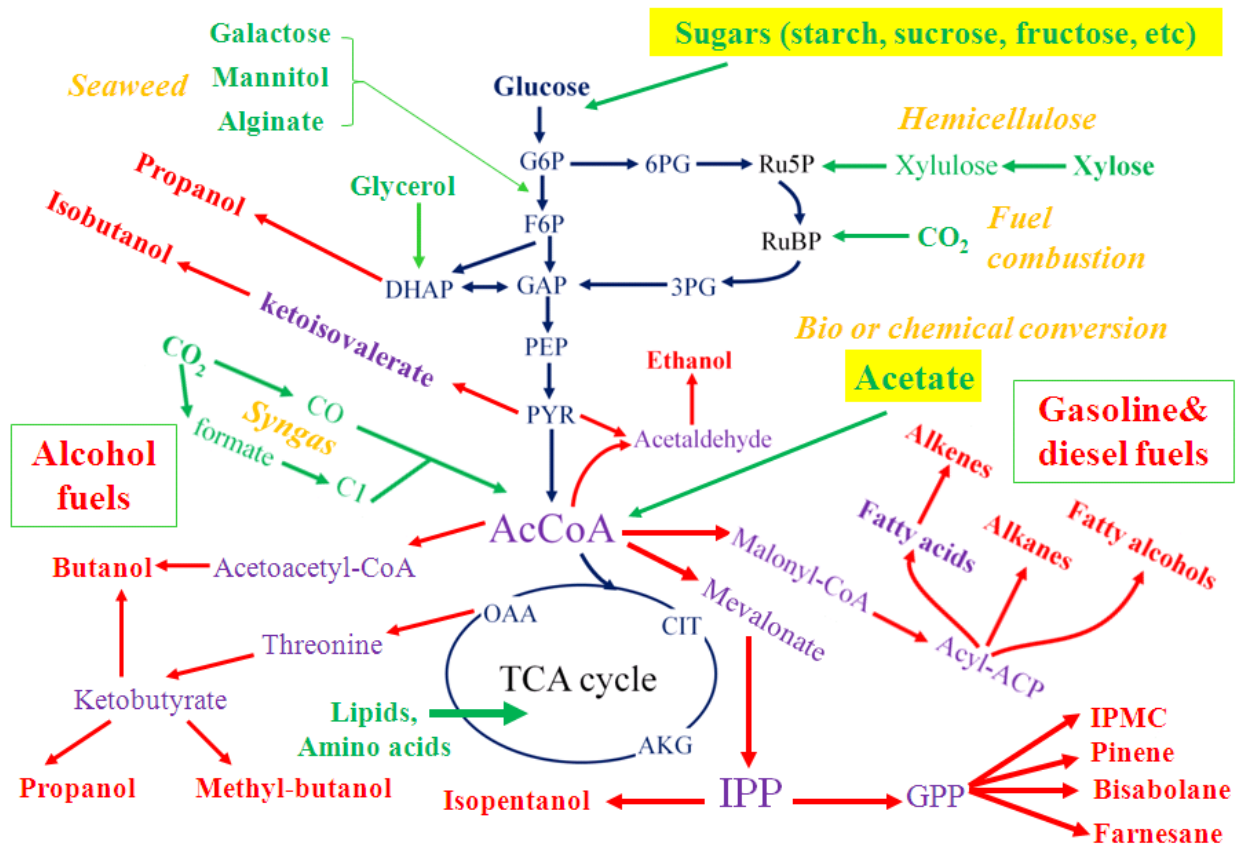


Figure 1.1: Metabolic engineering pathways for biofuel production

Ethanol is currently the most commercially successful biofuel and can be produced by yeast fermentations. Yeast efficiently converts sugar into ethanol and CO₂ via glycolysis pathway and pyruvate decarboxylase / alcohol dehydrogenase. Biofuels with properties similar to those of

gasoline and diesel fuel are being synthesized by microorganisms (Figure 1.1). Several engineered biofuel pathways are being examined. For example, engineered *Escherichia coli* can use the keto acid pathway and the Ehrlich pathway to produce higher alcohols (such as isobutanol), while the mevalonate pathway in yeast can be extended to synthesize branched and cyclic hydrocarbons (the biofuels with lower freezing point and higher energy content).

Table 1.1: Commonly employed microbes for biofuel production

Species	Substrates	Products	Features
<i>Saccharomyces cerevisiae</i>	Glucose, fructose, galactose, and others	Alcohols	Easy genetic manipulations, Crabtree effect
<i>Zymomonas mobilis</i>	Glucose, fructose, sucrose	Ethanol	High ethanol tolerance and yield
<i>Clostridium thermocellum</i>	Glucose, cellulose, cellobiose	Ethanol	Growth at high temperature, mixed fermentation pathways
<i>Clostridium acetobutylicum</i>	Glucose, xylose	Ethanol and butanol	Acetone, ethanol, and butanol fermentation
<i>Escherichia coli</i>	Glucose, xylose, glycerol, and others	Alcohols, diesels, and other biofuels	Easy genetic manipulations, fast growth
Cyanobacteria (e.g., <i>Synechocystis</i> 6803)	CO ₂	Alcohols, H ₂ , fatty acids	CO ₂ fixation
<i>Phanerochaete chrysosporium</i>	Glucose and lignin	cellulosic biomass pretreatment	Strong ability to degrade lignin
<i>Yarrowia lipolytica</i>	Glucose, acetate and fatty acids	Lipids	Oleaginous yeast that accumulates lipids

Finally, microbial metabolisms for biofuel production are very different across the species. *Saccharomyces cerevisiae* and *Escherichia coli* are microbial cell factories that are widely used in biofuel industrial because the two model species ferment sugar efficiently and are

also amenable to genetic modification and bioprocess scale up. Other microbial species, such as cyanobacteria, are also promising hosts for biofuel production because they can convert sunlight and CO₂ to biomass and products. The species diversity in metabolic features offers opportunity for synthesizing many different useful products from diverse carbon substrates. Table 1.1 shows several different microbial species that produce biofuels, either via the native biofuel pathway or via a metabolically engineered pathway.

1.3 Tools for genetic engineering of microbial hosts

Overexpression of native or heterologous genes is often achieved through plasmid based expression systems. Plasmids are naked DNA molecules that are capable of replication within the host. Plasmids are commonly used to carry genetic materials and transfer them to the microbial host. The gene expression is mainly controlled at the transcript level, i.e., by tuning with the promoter. The most widely used promoters are the *lac* and the hybrid promoters such as *tac*, *tic* and *trc*. These promoters can be induced under the presence of isopropyl-β-D-thiogalactopyranoside (IPTG). However, there is also research being done for the use of constitutive promoters for gene expression. Some of them rely on the use of natural promoters and others rely on random mutation of constitutive promoters. One of the goals of synthetic biology is to manipulate protein expression at the translation level and this can be achieved by modulation of the ribosomal binding site. Riboregulators have been developed to tune gene expression by RNA-RNA interactions. Another method by which gene targets can be overexpressed is through codon optimization⁴.

Plasmid based expression systems often suffer from unstable genetic performance. Chromosomal integration of target genes along with the promoter can be utilized to avoid this

problem. Often, overexpression of many genes would be required to achieve a desired yield. Novel approaches have been developed to control the coordinated expression of each gene. One such approach combines multiple genes into an operon under the control of a single promoter and the expression of each gene is controlled at posttranscriptional stage by tuning the intergenic regions. This method was applied for the coordinated expression of three genes of the heterologous mevalonate pathway in *E. coli* and resulted in an increase of the mevalonate production by sevenfold¹².

Knockout of competing pathways can redirect the flux of carbon towards the product of interest. Gene deletion is often achieved through homologous recombination and traditionally this is performed through plasmids containing a selectable marker flanked by DNA fragments of the target gene. Genes can be deleted in yeast by the use of a linear PCR fragment along with a short flanking region homologous to the target DNA. Gene deletions can also be performed using bacteriophage, and they depend on the FLP-FRP recombination to remove the marker after gene deletion. This method leaves a 68bp FRT scar on the chromosome for each deletion performed.

1.4 Cyanobacteria as a microbial cell factory

Direct capture of CO₂ for the synthesis of bioproducts is a more economical and environmental friendly approach that has received extensive studies recently. Cyanobacteria or blue-green algae are photoautotrophic prokaryotes and can fix CO₂ in the presence of sunlight. The photosynthetic efficiency of cyanobacteria is much higher than that of higher plants (10 – 20% in contrast to 0.5% in higher plants)¹. The transformability of some cyanobacteria species coupled with the availability of sequenced genomes allows us to perform complex genetic

engineering¹³. They generally have high growth rates as compared to green algae and plants. The diversity of metabolic capability in cyanobacteria lets them grow in highly saline environments as well as marginal lands and hence will not compete with land used for agriculture¹⁴. Among all cyanobacterial species, *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803) is one of the most extensively studied species since it was initially isolated from a freshwater lake in 1968. The entire genome, including four endogenous plasmids, was sequenced in 1996, and over 3000 genes have been annotated to date^{15, 16}. *Synechocystis* 6803 demonstrates versatile carbon metabolisms, growing under photoautotrophic, mixotrophic and heterotrophic conditions¹⁷. Additionally, biochemical similarities between the plant chloroplasts and *Synechocystis* 6803 make the latter an ideal system for studying the molecular mechanisms underlying stress responses and stress adaptation in higher plants¹⁸. More importantly, this species is naturally competent (homologous recombination at high frequency)¹⁹. The recent developments in synthetic biology have provided plenty of molecular biology tools to engineer *Synechocystis* 6803 as a photosynthetic host for the production of diverse types of chemicals.

Metabolic engineering has been applied for microbial ethanol production, including overexpression of genes to increase ethanol yield, disruption of genes to direct the carbon flux to ethanol and deletion of enzymes that can oxidize NADH. To directly convert CO₂ to biofuel, the cyanobacterium *Synechococcus* sp. PCC 7002 was engineered for the synthesis of ethanol²⁰. Recently, *pdc* and *adh* genes from *Zymomonas mobilis* were integrated into the chromosome of *Synechocystis* sp. PCC6803 under the control of the strong light driven *psbA2* promoter. An average yield of 5.2 mmol ethanol OD₇₃₀ unit⁻¹ litre⁻¹ day⁻¹ was achieved²¹. Algenol Biofuels Inc. claim that they can produce ethanol at a rate of 6000 gal/acre/year from an engineered cyanobacterial strain²².

Butanol is hydrophobic, has greater energy density, and a higher octane rating relative to ethanol. Therefore, butanol biosynthesis has received extensive studies. Acetone-butanol-ethanol (ABE) fermentation uses *Clostridium acetobutylicum* to produce *n*-butanol, but such process is restrained by relatively low production rate and generates large amount of byproducts. To overcome this problem, the *n*-butanol pathway derived from *Clostridium* was reconstructed in fast-growing *E. coli* or yeast strains^{23, 24}. Another novel alcohol synthesis approach is via non-fermentative pathway²⁵, where the amino acids biosynthesis pathways and Ehrlich pathway^{26, 27} were utilized to convert glucose to alcohols. Cyanobacteria have been explored for biofuel production (Figure 1.2). *Synechococcus elongatus* PCC 7942 was engineered to accumulate 450 mg/L of isobutanol in 6 days²⁸. *S. elongatus* 7942 was engineered with a modified CoA-dependent 1-butanol pathway and this strain accumulated 14.5 mg/L 1-butanol under anoxic condition²⁹. Long chain alcohols and hydrocarbons have ideal properties for combustion and are found to be either additives or major components of petroleum. *Synechocystis* 6803 and *Arabidopsis thaliana* were engineered with a heterologous fatty acyl-CoA reductase (FAR) for the production of fatty alcohols³⁰. Researchers at LS9 identified two key enzymes responsible for the production of alkanes in cyanobacteria: an acyl-CoA carboxylase and an aldehyde decarbonylase³¹. This discovery opens up many possibilities for engineering cyanobacteria for alkane production.

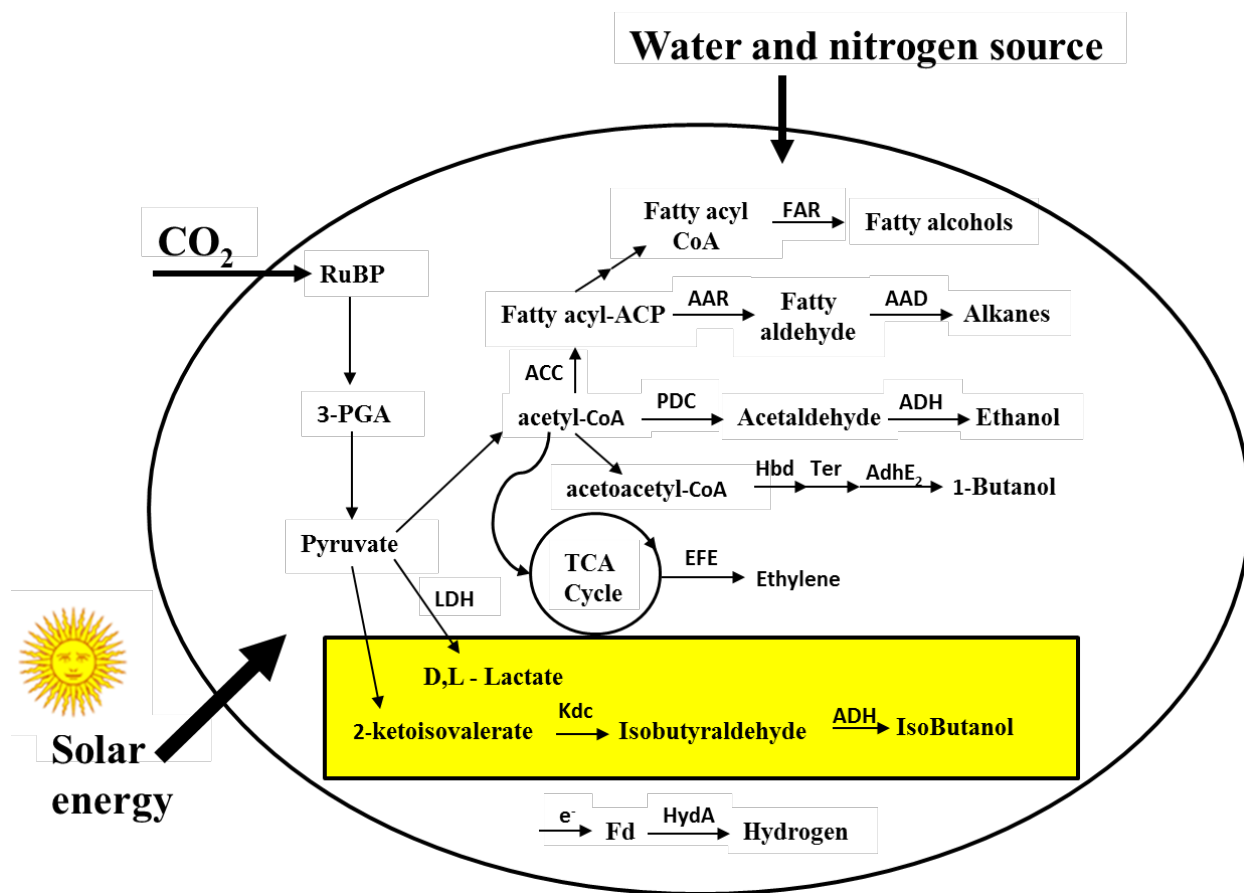


Figure 1.2: Schematic representation of engineered biochemical pathways for the production of biofuels in cyanobacteria. Abbreviations: RuBP, Ribulose-1,5-bisphosphate; 3-PGA; 3-phosphoglycerate; Kdc, ketoacid decarboxylase; ADH, alcohol dehydrogenase; PDC, pyruvate decarboxylase; ACC, acetyl-CoA carboxylase; AAR, acyl-ACP reductase; AAD, aldehyde decarbonylase; FAR, fatty acyl-CoA reductase; Ter, trans-2-enoyl-CoA reductase; Hbd, 3-hydroxybutyryl-CoA dehydrogenase; AdhE2, aldehyde/alcohol dehydrogenase; LDH, lactate dehydrogenase; EFE, Ethylene formation enzyme. Highlighted reactions indicate the pathway that will be focused in this study.

Finally, many of the cyanobacterial strains have native hydrogenases that can evolve hydrogen under anoxic conditions. Though the theoretical efficiency for hydrogen production is predicted to be high, the efficiency in which wild type cyanobacterial strains produce hydrogen was observed to be very low ($< 0.1\%$)¹³. *Synechococcus elongatus* sp. 7942 was engineered with hydrogenase from *Clostridium acetobutylicum* and was demonstrated that the

heterogeneous hydrogenase can support hydrogen evolution at a rate >500 times than that of endogenous hydrogenase under anoxic conditions³².

1.5 Modeling and Systems Analysis

Mathematical kinetic models can be helpful in interpreting experimental data, in understanding quantitative functions of biological systems, and in predicting metabolic performances³³. Arnold Fredrickson introduced the terms “segregated” and “structured” to categorize most kinetic models for biological systems. The term “segregated” was used to take into account the presence of heterogeneous individuals in a cell population explicitly (For example: a model that would take into account the different age groups of cell that would be present in a cell population). The “structured” kinetic model was used to define formulation of cell systems as composed of multiple biomass components. The group “Unsegregated and unstructured” is the most idealized case which considers the cell population as one component solute and most of the kinetic models will fall in this category (e.g., Michaelis-Menten kinetics: $V=V_m \times S / (K_m + S)$ ³⁴).

On the other hand, metabolic fluxes do not consider kinetic behavior of microorganisms, but they provide the ratios in which each pathway is engaged in cellular functions. Fluxomics in an organism were first studied using *in silico* analysis known as Flux Balance Analysis (FBA). FBA uses the stoichiometry of metabolic reactions along with a set of constraints³⁴. The total number of reactions and constraints is often less than the number of variables (Flux) to be calculated and hence the system is underdetermined. This necessitates the use of an objective function to calculate the set of theoretical fluxes. Maximization of biomass is the objective function employed generally as all species evolve themselves to multiply more in their

environment. ^{13}C -MFA computes the overall pathway activities in an organism by utilizing the isotopic labeling approach and is valid only at isotopic and metabolic steady state. ^{13}C -MFA is performed by feeding the microbes with a ^{13}C labeled carbon source followed by measurement of the enrichment pattern of the metabolites. The isotopic labeling pattern of all the metabolites is then fed to an algorithm to generate the intracellular fluxes³⁴. Though both the methods of FBA and ^{13}C -MFA utilize the overall metabolic network and use the assumption of metabolic steady state, they have two different purposes. FBA gives an optimal flux distribution to achieve a desired performance whereas ^{13}C -MFA quantifies the *in vivo* operation of a cell. The two techniques complement each other and can be utilized to locate bottlenecks in metabolic pathways for the synthesis of a desired product (Figure 1.3).

Finally, the current flux analysis disregards the dynamic metabolic behavior of a biological system. This avoids the difficulties in solving large-scale kinetic models and performing time-dependent experimental measurements. However, many biological systems cannot maintain a metabolic (or isotopic) steady state during the entire cultivation process. The flux modeling for dynamic metabolite concentrations or isotopic labeling requires the innovative approaches to link kinetic model to metabolic flux analysis.

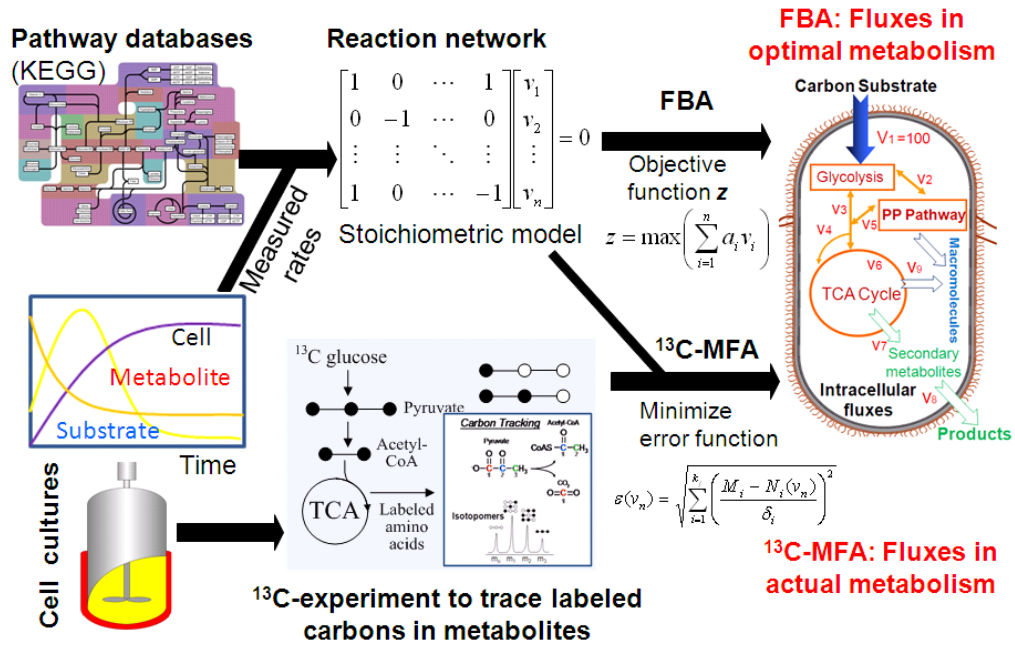


Figure 1.3: Metabolic network modeling and analysis. FBA profiles the "optimal" metabolism for the desired performance (can be genome-scale); ^{13}C -MFA measures in vivo operation of the central metabolic network (< 100 reactions).

Chapter 2: Statistics-based model for prediction of chemical biosynthesis yield from *Saccharomyces cerevisiae*

This chapter has been reproduced from the following publication:

Varman, A.M., Xiao, Y., Leonard, E. & Tang, Y. Statistics-based model for prediction of chemical biosynthesis yield from *Saccharomyces cerevisiae*. *Microbial Cell Factories* **10**, 45 (2011).

Abstract

Background

The robustness of *Saccharomyces cerevisiae* in facilitating industrial-scale production of ethanol extends its utilization as a platform to synthesize other metabolites, both native and of heterologous origins. Metabolic engineering strategies, typically via pathway overexpression and deletion, continue to play a key role for optimizing the conversion efficiency of substrates into the desired products. However, chemical production titer or yield remains difficult to predict based on reaction stoichiometry and mass balance. We sampled a large space of data of chemical production from *S. cerevisiae*, and developed a statistics-based model to calculate production yield using input variables that represent the number of enzymatic steps in the key biosynthetic pathway of interest, metabolic modifications, cultivation modes, nutrition and oxygen availability.

Results

Based on the production data of about 40 chemicals produced from *S. cerevisiae*, metabolic engineering methods, metabolite supplementation, and fermentation conditions described therein, we generated mathematical models with numerical and categorical variables to predict

production yield. Statistically, the models showed that: 1. Chemical production from central metabolic precursors decreased exponentially with increasing number of enzymatic steps for biosynthesis (>30% loss of yield per enzymatic step, P-value=0); 2. Categorical variables of gene overexpression and knockout improved product yield by 2~4 folds (P-value<0.1); 3. Addition of notable amount of intermediate precursors or nutrients improved product yield by over five folds (P-value<0.05); 4. Performing the cultivation in a bioreactor enhanced the yield of product by three folds (P-value<0.05); 5. Contribution of oxygen to product yield was not statistically significant. Yield calculations for various chemicals using the linear model were in fairly good agreement with the experimental values. The model generally underestimated the ethanol production as compared to other chemicals, which supported the notion that the metabolism of *Saccharomyces cerevisiae* has historically evolved for robust alcohol fermentation.

Conclusions

We generated simple mathematical models for first-order approximation of chemical production yield from *S. cerevisiae*. These linear models provide empirical insights to the effects of strain engineering and cultivation conditions toward biosynthetic efficiency. These models may not only provide guidelines for metabolic engineers to synthesize desired products, but also be useful to compare the biosynthesis performance among different research papers.

2.1 Background

Producing small-molecule chemicals from microbial biocatalysts offers several advantages. Unlike conventional chemical synthesis which are heavily dependent on petroleum-derived substrates, microbes are able to use renewable materials to synthesize many commodity

chemicals and fuels ⁷ (Figure 2.1). Due to its scalability, microorganisms are also suitable platforms to synthesize pharmaceutical molecules that are conventionally produced from extracting large amounts of natural resources. Among many industrial microorganisms, the baker's yeast, i.e., *S. cerevisiae* continues to emerge as a preferred production platform ³⁵. *S. cerevisiae* is typically known for its robustness in fermenting sugars into alcohol. In the recent past, it has also gained importance as a heterologous platform to synthesize many precursors of commodity chemicals and pharmaceuticals ⁷. In general, chemical production using whole-cell biocatalysts are achieved by genetic engineering to extend the substrate range of an existing biosynthetic pathway or to introduce new biosynthetic pathways (either derived from other organisms, or completely novel). Rational metabolic engineering approaches then analyze the cellular metabolism and improve production titer by overexpressing rate-limiting enzymes or deleting competing pathways. In general, the actual yield of chemical production is not easily predicted due to the complexity of biological systems and dependency of cultivation conditions. Biological complexities not only include intrinsic properties (such as enzyme kinetics and substrate specificity), but also include enzyme compartmentalization, intracellular signaling, and metabolite transport between eukaryotic cell organelles. Therefore, strain engineering requires multiple rounds of trial-and-error experiments to perform the optimum combination of genetic manipulations. In the present work, we sought to develop mathematical models that could provide *a priori* estimation of chemical production yield from engineered *S. cerevisiae* when given a set of parameters, namely the number of steps in the biosynthetic pathway of interest, genetic modifications, cultivation conditions, and nutrient and oxygen availability. The coefficients of these parameters were obtained from the regression of the yields and production conditions reported by recent literatures. Such model predicted the empirical yields that were

lower than the theoretical productivities under “ideal” conditions. The model results could give metabolic engineers guidelines for increasing desired products and for reducing futile attempts.

2.2 Model development

The model defined several important parameters that influenced the efficiency of chemical production from microbial hosts. The first group of parameters accounted for the number of enzymatic steps in the biosynthetic pathway of interest since it had been shown that this parameter was often inversely correlated with microbial product yield⁹. To enumerate the number of enzymatic steps, we introduced two numerical variables in our model, i.e. PRI and SEC. The variable PRI specified the number of enzymatic steps in primary metabolism (Figure 2.1), e.g. glycolysis that is required to convert sugar (glucose or galactose) to pyruvate. The variable SEC specified the number of enzymatic steps in the subsequent pathway (typically belongs to secondary metabolism), which catalyzed the conversion of central carbon intermediate into the final product of interest. The next group of variables was to capture the effects of genetic modification. Various genetic strategies have been used to implement metabolic engineering^{4, 34}. For example, promoters with different strength influence production level. However, for the sake of simplifying our model, variations of genetic components used in metabolic engineering strategies were lumped into two ordinal variables, i.e. OVE, and KNO. OVE signified the introduction of multiple copies of genes of native or heterologous origin for the purpose of improving production level. KNO signifies the alteration of branch pathways that might compete with the pathway of interest^{5, 36}. We further sub-categorized OVE based on the number of modified genes into OVE_{C1} (without “pushing” pathway flux), OVE_{C2} (enhancing 1~2 enzyme activities), and OVE_{C3} (improving a number of key enzyme functions). KNO was also

categorized by KNO_{C1} and KNO_{C2} (i.e., without knockout or with knockout, respectively). Table 2.1 explained the specifications for each sub-category.

The yield of metabolite production is also a function of cultivation conditions and nutrient availability. For instance, production of metabolites from a bioreactor is often higher than a shaking flask, due to the increased efficiency of mass transfer of oxygen, substrates, and nutrients. Moreover, culture acidification that often generates cytotoxicity and maintenance burden to the microbial hosts can be mitigated in a bioreactor by automated pH control. Based on these basic properties, we introduced the variable CUL to represent the general property of a cultivation condition. We also introduced the variable OXY and NUT to capture the effects of oxygen availability and nutrient supplementation, respectively³⁷⁻³⁹. Moreover, the variable INT captured the effect of addition of a secondary carbon source which served as a precursor or an intermediate metabolite of the pathway of interest.

Several assumptions were made to simplify our model development. A) Yield calculation was based on the conversion of major carbon substrate to final product if multiple nutrient sources were supplemented (e.g., yeast extract was not treated as the carbon source). B) We calculated the yields based on two factors: initially added carbon substrate in the culture and final measured product. We neglected the unused carbon substrate that remained in the end of the production. C) To calculate enzymatic steps from the carbon source, the model only considered the key route from the major substrate (mostly glucose) to the final products (enzyme steps for co-factors or ATPs synthesis were neglected). D) For product synthesis promoted by the addition of an intermediate, we had no means of differentiating the carbons derived from added precursor or from the carbon substrate (i.e., glucose). To account for the contribution from both carbon

sources, the yield calculation was assumed to be an arithmetic mean of the two yields (One yield was based on substrate, e.g., glucose, and the other yield was estimated from the intermediates). Meanwhile, the number of primary steps or secondary steps were also assumed as an arithmetic mean of two data sets (one variable was counted from substrate; the other variable was counted from the intermediate).

Biochemical systems theory³⁴ states that reaction rates (v_i) can be described by a general power law expression of the type:

$$v_i = \alpha_i \prod_j X_j^{g_{ij}} \quad (2.1)$$

Where X_j represents the system variables and the parameters α_i , g_{ij} are the constants. Equation (2.1) yields a linear form in logarithmic coordinates. Based on similar assumptions, our model for yield prediction used system variables (i.e., numerical or categorical variables related to yeast biosynthesis) to describe the relative carbon flux to the final products.

$$\log_{10} Y = \beta_0 + \beta_{\text{PRI}}\text{PRI} + \beta_{\text{SEC}}\text{SEC} + \beta_{\text{OVE,C2}}\text{OVE}_{\text{C2}} + \beta_{\text{OVE,C3}}\text{OVE}_{\text{C3}} + \beta_{\text{KNO,C2}}\text{KNO}_{\text{C2}} + \beta_{\text{NUT,C2}}\text{NUT}_{\text{C2}} + \beta_{\text{INT,C2}}\text{INT}_{\text{C2}} + \beta_{\text{CUL,C2}}\text{CUL}_{\text{C2}} + \beta_{\text{OXY,C2}}\text{OXY}_{\text{C2}} \quad (2.2)$$

In Equation 2.2, $\log_{10} Y$ was the dependent variable which represented production yield (mol C in product/mol C in primary substrate), given each independent variables β_i ⁴⁰. We defined β_0 as the intercept in Equation 2.2, which represented the combined contribution of Category 1 of all ordinal variables. β_0 was defined as:

$$\beta_0 = \beta_{\text{OVE,C1}} + \beta_{\text{KNO,C1}} + \beta_{\text{NUT,C1}} + \beta_{\text{INT,C1}} + \beta_{\text{CUL,C1}} + \beta_{\text{OXY,C1}} \quad (2.3)$$

The ordinal variables (using a binary system) were assigned a value of one if and only if the condition fitted the category in Table 2.1. Otherwise, the ordinal variables were assigned a value of 0⁴¹. (2) To acquire the coefficients in Equation 2.2 and 2.3, we compiled data from ~40 publications which described the production of chemicals by *S. cerevisiae* under various experimental conditions. Table 2.2 summarized the categories assigned to these experimental conditions and the yield of product from our best judgment. Using these data, we performed regression analysis to fit the model via the software package R⁴² to find the regression coefficients and P-values. For this study, a variable was statistically significant (90%) if its P-value was below 0.1.

2.3 Results and discussion

We constructed simple models which linked several numerical and ordinal variables that affected the yield of chemical production from *S. cerevisiae*. These ordinal variables consisted of the number of modified genes or pathways (OVE), the number of gene knockouts in known competitive pathways (KNO), nutrient source (NUT), intermediate (INT), cultivation mode (CUL), and oxygen availability (OXY). We described the yield of chemical production as the summation of these independent variables in Equation 2.2. We fitted Equation 2.2 and determined the coefficients of the variables using linear regression analysis of ~40 compounds. Although multiple data of production yields were often reported in each literature, the model only considered the best yield under a denoted experimental condition. Then, all experimental conditions were categorized by numerical and ordinal variables. The linear regression coefficients obtained for Equation 2.2 were given in Equation 2.4, such that:

$$\log Y = -1.53 - 0.01 \text{ PRI} - 0.19 \text{ SEC} + 0.007 \text{ OVE}_{\text{C2}} + 0.52 \text{ OVE}_{\text{C3}} + 0.31 \text{ KNO}_{\text{C2}} + 0.73 \text{ NUT}_{\text{C2}} + 0.77 \text{ INT}_{\text{C2}} + 0.51 \text{ CUL}_{\text{C2}} + 0.27 \text{ OXY}_{\text{C2}} \quad (2.4)$$

The accuracy of obtained coefficients in Equation 2.4 was evaluated based on R^2 and the P-value. Here, we used a P-value of 0.1 as the limit below which the result was considered significant⁴³. Out of the eight variables specified in our model, SEC, OVE, KNO, NUT, INT and CUL had P-value of less than 0.1. The summary of the P-value of each variable was listed in Table 2.3. Figure 2.2A showed a plot of the production yields obtained experimentally and those obtained from model prediction for the corresponding conditions. The correlation of this model to the dataset had an R^2 value of 0.55, which reflected the moderate discrepancy between reported yields and the model-predicted yields. Figure 2.2B plotted the residuals of model fitting. The residuals appeared to scatter around zero randomly, so the linear model was proper to describe the experimental data.

Interestingly, the number of enzymes in the primary pathway (PRI) did not significantly affect production yield (P-value = 0.76) (Table 2.3). This suggested that rate-limiting steps to increase chemical production flux often lay in the downstream pathway of central metabolism. The coefficient of SEC was negative. This suggested that the length of a pathway downstream of central metabolism negatively affected production yield. Specifically, addition of a new enzymatic step in a secondary metabolic pathway reduced product yield by 36% (for numerical variable SEC: $10^{\beta_{\text{SEC}}} = 10^{-0.19} = 64\%$). A good demonstration of the effect of pathway length on product yield was found in the case of naringenin production⁴⁴. With the following inputs of variables PRI = 10 (Galactose to PEP), SEC = 14 (i.e., 10 steps from PEP to phenylalanine; 4

steps from phenylalanine to flavanone), KNO = INT = CUL = OXY = category 1, NUT = Category 2; OVE = Category 3; the model calculated:

Yield = $10^{-1.53 - (0.01 \times 10) + (-0.19 \times 14) + 0.52 + 0.73} = 0.0009$ (The reported experimental production yield was 0.00058). In most cases, our model-predicted yields were within the range of one order of magnitude compared to the experimental values.

Since the number of steps in central metabolism (PRI) did not significantly affect production yield, we computed another set of regression coefficients for Equation 2 without the variable PRI, to yield a simplified form Equation 2.5.

$$\log Y = -1.60 - 0.19 \text{ SEC} + 0.0003 \text{ OVE}_{C2} + 0.50 \text{ OVE}_{C3} + 0.31 \text{ KNO}_{C2} + 0.73 \text{ NUT}_{C2} + 0.82 \text{ INT}_{C2} + 0.51 \text{ CUL}_{C2} + 0.28 \text{ OXY}_{C2} \quad (2.5)$$

As shown in Table 2.3, regression using Equation 2 with the exclusion of the variable PRI did not change the R^2 value. This result indicated that the number of enzymatic steps in primary metabolism did not significantly affect product yield. Presumably, fluxes in central metabolic pathways were typically high and robust⁴⁵, when compared to those downstream secondary pathways. It has been demonstrated recently that production of chemicals was significantly improved, only when the capacity of a downstream pathway was increased⁴⁶.

Metabolic engineering typically involves pathway modification to shift metabolic fluxes into a desired product or to permit the use of an alternative carbon source⁴⁷. We defined the variable OVE, and KNO in Equation 2.2 to capture the effect of pathway overexpression, and deletion, respectively. The regression of experimental data using Equation 2.2 showed that the coefficients of OVE_{C2} and OVE_{C3} had positive values (Table 2.3). The model successfully captured the

contribution of both pathway overexpression and gene deletions to increase product yield in *S. cerevisiae*. The high P-value of OVE_{C2} (0.98) indicated that statistically, the overexpression of a small number of genes (1-2) was uncertain to improve production yield. However, the coefficient of OVE_{C3} (=0.52; P-value=0.07) indicated the effectiveness of multiple gene modification to resolve the bottleneck steps. This observation is consistent to the fact that metabolic fluxes generally do not sensitively respond to changes of single enzyme activity, but are controlled by all key enzymes along the biosynthesis pathway. On the other hand, the regression coefficients of KNO_{C2} had positive value (=0.31, P-value = 0.08), and thus the removal of competitive pathways could be effective to increase production yield.

It is a general knowledge that bioprocess conditions affect cellular viability and product yield. Our model suggested fermentation using a well-controlled bioreactor improved production yield by 3.2 times ($CUL_{C2}:10^{\beta_{CUL,C2}} = 10^{0.51}$). The model further suggested that fermentation under anaerobic or microaerobic condition could enhance yield compared to aerobic fermentation. However, such enhancement was not statistically significant (P-value = 0.32). This observation could be explained by the fact that *S. cerevisiae* produced fermentative products (ethanol and glycerol) (Crabtree effect)^{48, 49} under aerobic and glucose-sufficient medium. Therefore, aerobic metabolism in *S. cerevisiae* could operate similarly to metabolism under oxygen-limited condition. The coefficient for the variable INT was 0.77, which represented that the supplementation of a precursor metabolite translated to an approximately six fold increase of the product yield (P-value = 0.02). Similarly, the addition of nutrients (such as yeast extract) also significantly increased production yield (the coefficient of NUT_{C2} was 0.73). The contributions of INT and NUT to product formation indicated that intermediates/nutrients provided building blocks or energy sources that reduced the rate-limiting steps in biosynthetic pathways.

We used Equation 2.2 to compute the production yield of chemicals according to the specifications listed in Table 2.2. We observed that, for ethanol production, the experimental values were generally higher than the empirical model predictions. In reality, the reported maximum ethanol yield could reach 0.5 mol C-ethanol / mol C-glucose⁵⁰, which could be several folds higher than model predictions. To mitigate this discrepancy, we re-categorized the ethanol synthesis pathway as the primary pathway to generate Equation 2.6.

$$\log Y = -1.73 + 0.003 \text{ PRI} - 0.19 \text{ SEC} + 0.05 \text{ OVE}_{\text{C2}} + 0.56 \text{ OVE}_{\text{C3}} + 0.37 \text{ KNO}_{\text{C2}} + 0.71 \text{ NUT}_{\text{C2}} + 0.86 \text{ INT}_{\text{C2}} + 0.51 \text{ CUL}_{\text{C2}} + 0.12 \text{ OXY}_{\text{C2}} \quad (2.6)$$

Regression of the data using Equation 2.6 improved the R^2 value from 0.55 to 0.58, demonstrating that ethanol could be better assumed as a central metabolite for *S. cerevisiae*. Using Equation 2.6, we predicted ethanol production based on a recent reference⁵¹ by specifying PRI = 11, SEC = 1 (cellulose degradation step), OVE = C3, KNO = C1; NUT = C2, INT = C1, CUL = C1, and OXY = C2. The ethanol production yield calculated by Equation 2.6 was 0.31. This value was in good agreement with the reported values of ~0.4⁵¹.

2.4 Model applications and limitations

The main application of the model is to predict the biosynthesis yield from *S. cerevisiae*. The model were validated by “unseen data” (Figure 2.2C) from some randomly selected new publications (2010~2011). The model predicted the yields based on the reported experimental conditions described by these papers⁵²⁻⁵⁶. Most yield data were close to model predictions. The predictive power of the model was consistent with the model quality described in Table 2.3.

Furthermore, the model can reveal the metabolic features of *S. cerevisiae*. For example, the modified model Equation 2.6 showed that it was better to treat ethanol pathway as the primary routes in cell metabolism, because of the strong ability for ethanol fermentation by yeast, possibly due to long-term process for selecting yeast as alcohol producer through human history. The model can also be useful for comparing the productivity among other yeast species (Figure 2.3). For example, riboflavin producer, *Candida famata*, exhibits a high riboflavin productivity (2~3 order of magnitude higher than model prediction)⁵⁷. *Pichia pastoris*, a common species for protein expression, shows high S-adenosyl-L-methionine productivity if a large amount of the intermediate methionine was repeatedly added in the medium⁵⁸. Besides, *Pichia stipitis* also has high yields of L-lactic acid and ethanol from glucose and xylose⁵⁹. Figure 2.3 demonstrated that some yeast species were able to explore their native pathways for biosynthesis of certain products with extraordinary efficiency (better than *S. cerevisiae*), therefore, these yeast species may be alternative hosts for certain biotechnology applications.

The accuracy of the model predictions for some products could be poor due to several limitations during model development. First, the category was a rough estimation of experimental conditions especially for variables related to gene modifications (OVE and KNO), and the yields could be very different even in the same category. Second, some products, despite large synthesis rates, were either not very stable or difficult to accumulate in a large quantity due to consumptions by downstream pathways or product degradations (e.g., Glycerol 3-phosphate⁶⁰). Their yields could be significantly lower than model predictions even though the actual flux to the product was high. Third, the coefficient β_{SEC} from model regression could not account for the big variances of biosynthesis efficiency or potentially feedback inhibitions in secondary pathways. For example, butanol synthesis is significantly improved via non-fermentative amino acid pathways compared

to traditional acetyl-CoA routes ⁶¹, because amino acid synthesis pathways in microorganisms are more effective than other heterogeneous pathways. Fourth, because of limited information from the references, the yield calculation could not precisely include the CO₂ fixation (e.g., overexpression of the native carboxylase pathway: pyruvate + CO₂ → oxaloacetate) ⁶² or the nutrients utilization in the rich medium. Fifth, the model neglected enzyme steps related to energy metabolism (such as ATP and NADPH synthesis), while cofactor imbalance can also affect the product yields.

2.5 Comparison to the previously published *E. coli* model ⁶³

Recently, we have constructed the *E.coli* model using same modeling approach. Compared to the *E.coli* model, *S. cerevisiae* shows several differences: 1. Oxygen conditions made a more significant impact on biosynthesis yield in *E.coli* than that in *S. cerevisiae*; 2. The genetic modification in *E.coli* had higher uncertainty for metabolic outcomes; 3. For metabolic pathways from precursors to final products, loss of yield per biosynthesis step (~30%) in *S. cerevisiae* is higher than that in *E.coli* (10~20%). Interestingly, *E. coli* model states that primary metabolism influences product yield (a relatively small P-value of 0.06) which indicates the balance of precursor production from central metabolism is also an important consideration for metabolic engineering of *E.coli*. For example, it has been demonstrated that lycopene production with *E. coli* was enhanced by redirecting the carbon flux from pyruvate to G3P ⁶⁴, but feeding other central metabolite precursors (such as pyruvate) could not improve lycopene production. On the other hand, the *S. cerevisiae* model indicates that it is less likely that the number of steps in central metabolism play a bottleneck role in the production of metabolites derived from it, while the bottlenecks are more likely in the secondary pathways (from central precursors to the

final product). Therefore, the metabolic strategies should focus on the secondary pathways to have a better chance for increasing final yield. Although modification of central metabolism may affect microbial physiologies, a few studies indicate the robustness of the central metabolism in *S. cerevisiae* because of its importance to cell vitality. For example, *S. cerevisiae* may maintain central metabolic fluxes via gene duplication and alternative pathways under different environmental and physiological conditions^{45, 65}. Therefore, the inflexibility of central pathways in *S. cerevisiae* is likely to render metabolic engineering strategies ineffective when targeting enzymes in central metabolism. In general, the unique metabolic features of yeast and bacteria can be of important consideration when choosing a production host.

2.6 Conclusions

Although *S. cerevisiae* has been widely used as a robust industrial organism for metabolic engineering applications, many metabolic features of this organism for biosynthesis under various conditions remain unknown. In this study, the statistic model for yeast biosynthesis permits *a priori* calculation of the final product yield achievable by current biotechnology. Unlike other *in silico* models based on mass balance or thermodynamics (such as FBA model)⁶⁶,⁶⁷, our model is based on a statistical analysis of published data using numerical and ordinal variables (categorized experimental conditions). The model has three applications. 1. The yield prediction takes into account the genetic design of the microbial host system and the “suboptimal” conditions under which the fermentation process occurs. 2. The model may identify effective metabolic strategies and at the same time, quantitatively provide the degree of uncertainty (i.e., possibility for failure). For example, statistical analysis shows that, for *S. cerevisiae*, metabolic bottlenecks may be more likely to be in the secondary metabolic pathways

rather than primary pathways, and thus it can narrow down the genetic targets and avoid futile work. 3. This model may be used to qualitatively benchmark yields of different engineered production platforms.

Acknowledgments

This study was supported by a grant from the National Science Foundation (MCB0954016). The authors also thank Bing Wu, Xueyang Feng, Peter Colletti and Yogesh Goyal for helping with data collection.

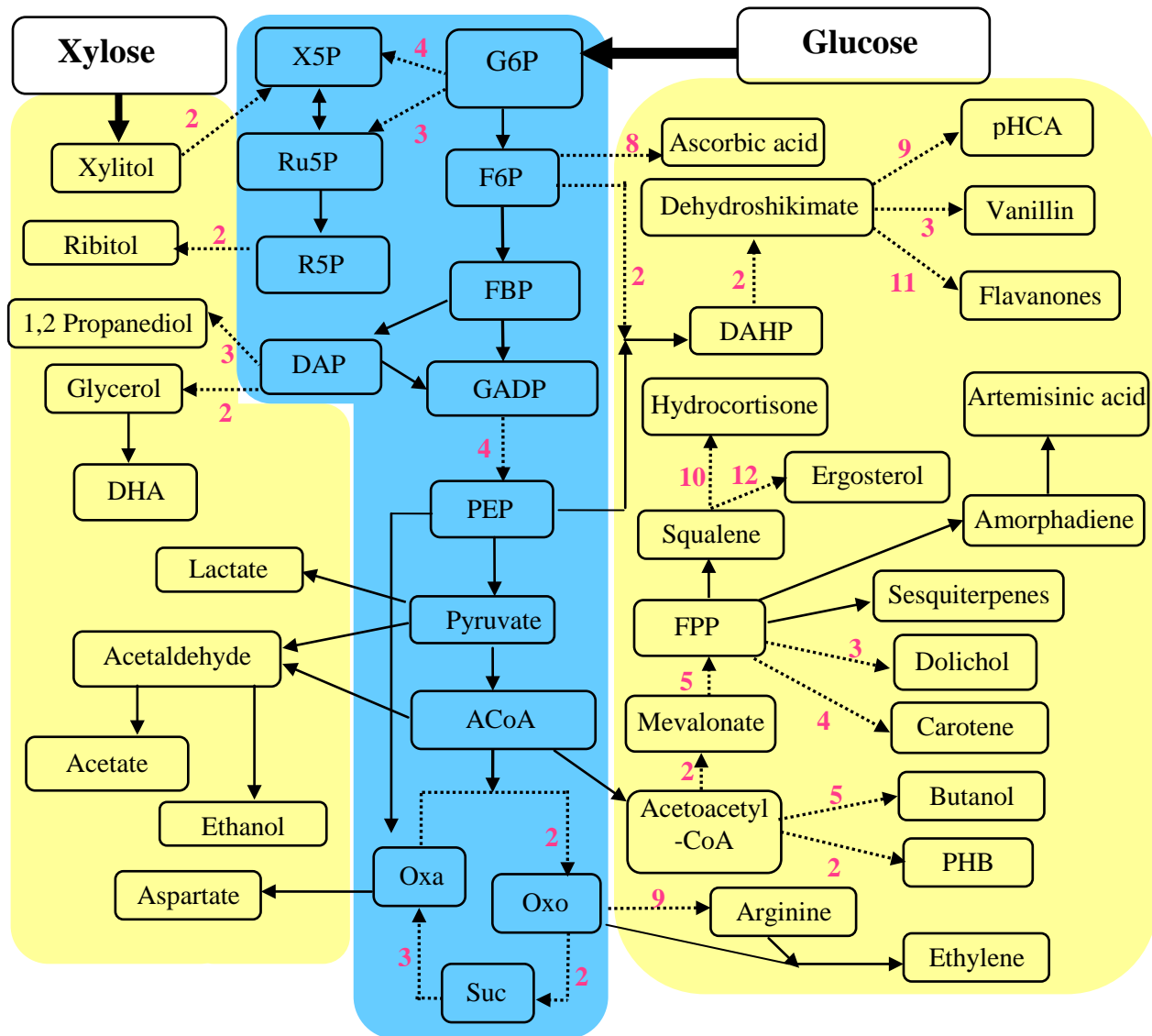


Figure 2.1: Metabolic pathway for the biosynthesis of major products. The blue box represents central metabolism and the yellow box represents secondary metabolism. Solid arrows signify single step reaction and dotted arrows signify multiple steps. **Abbreviations:** ACoA – Acetyl-CoA; DAP – Dihydroxyacetone-Phosphate; DAHP – 3-Deoxy-D-Arabino-Heptulose-7-Phosphate; DHA – Dihydroxyacetone; F6P - Fructose-6-Phosphate; FBP – Fructose 1,6-bisphosphate; G6P – Glucose-6-Phosphate; GADP – Glyceraldehyde-3-Phosphate; Oxa – Oxaloacetate; Oxo – 2-Oxoglutarate; PEP – Phosphoenolpyruvate; PHB – Poly[(R)-3-hydroxybutyrate]; pHCA – p-Hydroxycinnamic acid; R5P – Ribose-5-Phosphate; Ru5P – Ribulose-5-Phosphate; Suc – Succinate; X5P – Xylulose-5- Phosphate.

Figure 2.2 A

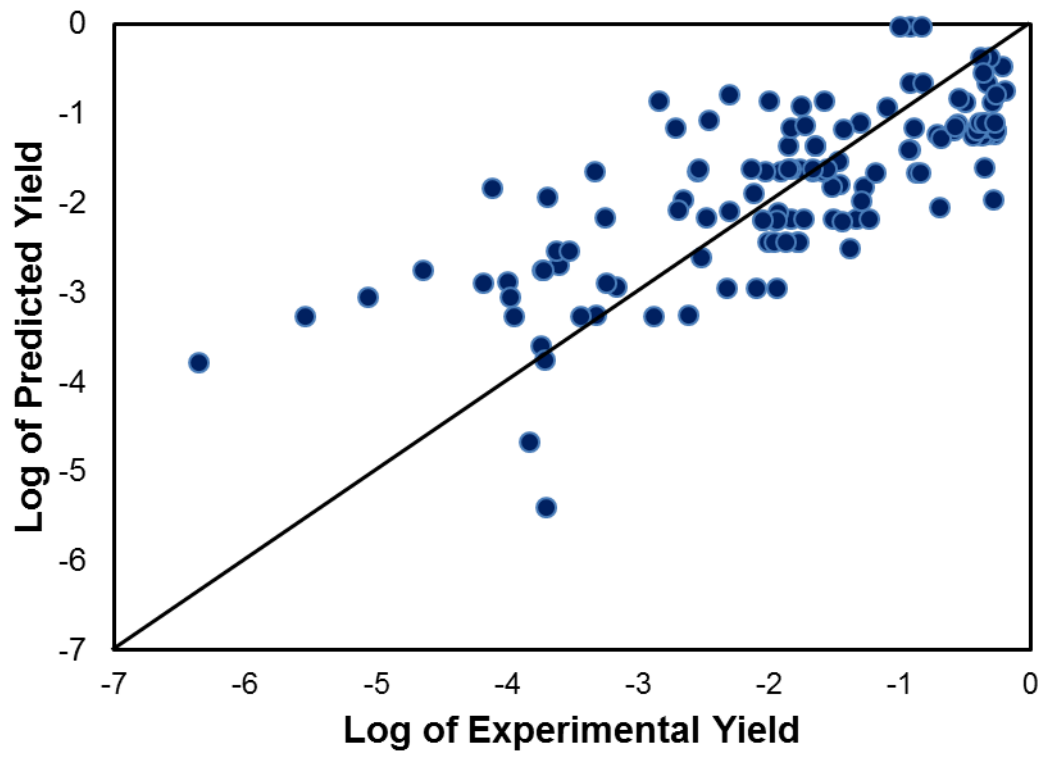


Figure 2.2 B

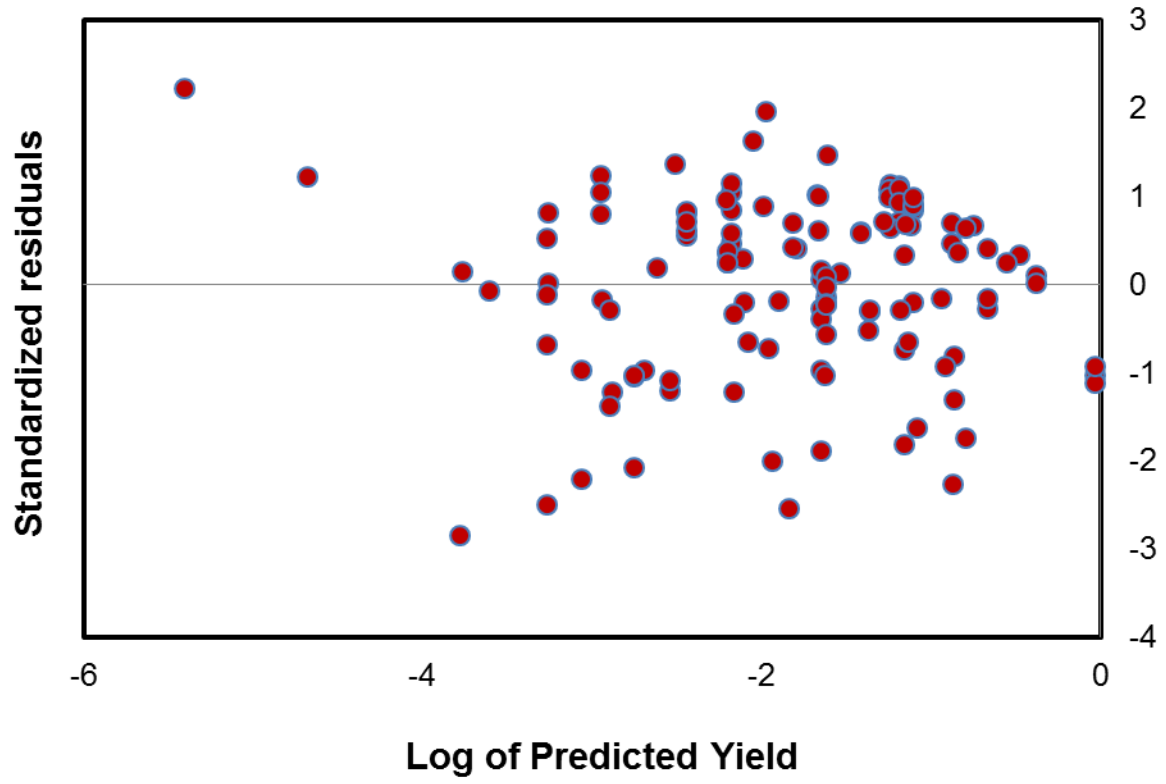


Figure 2.2 C

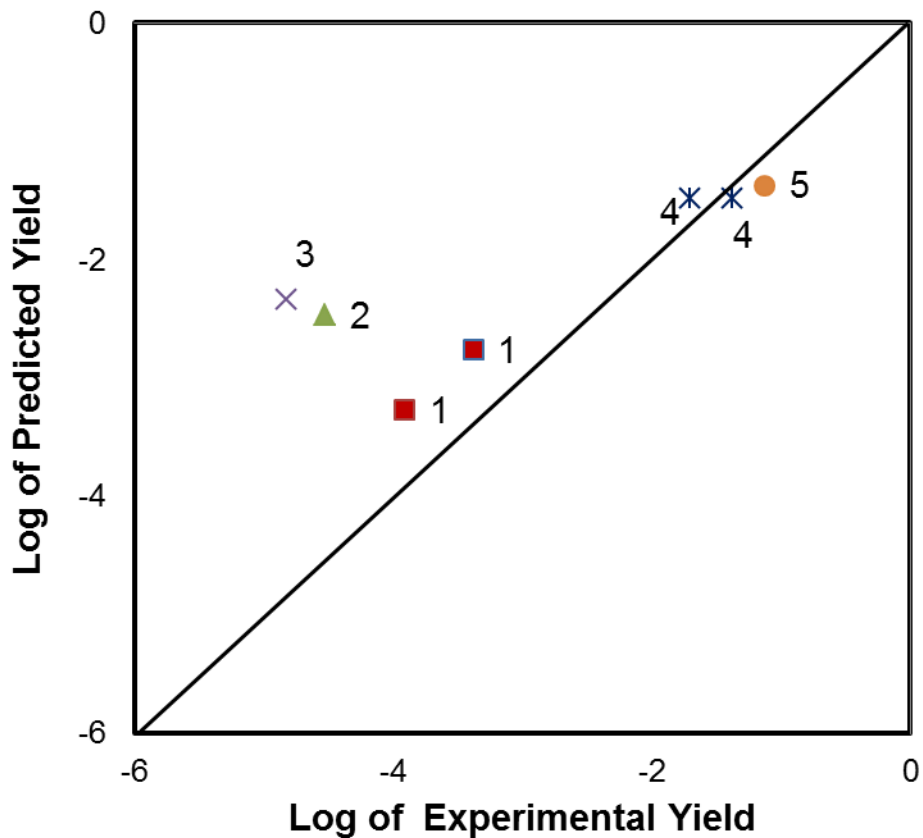


Figure 2.2: Model results. A) Plot of the actual logarithmic yields against the logarithmic yields generated by the regression model. The line drawn as diagonal to the plot is one-to-one and passes through the origin. The data points have an R² value of 0.55. B) Plot of residuals against fitted values. C) Model validation using newly published data (2010~2011) 1 - β -amyrin⁵²; 2 - ascorbic acid⁵³; 3 - monoterpene⁵⁴; 4 - vanillin⁵⁵; 5 - succinic acid⁵⁶.

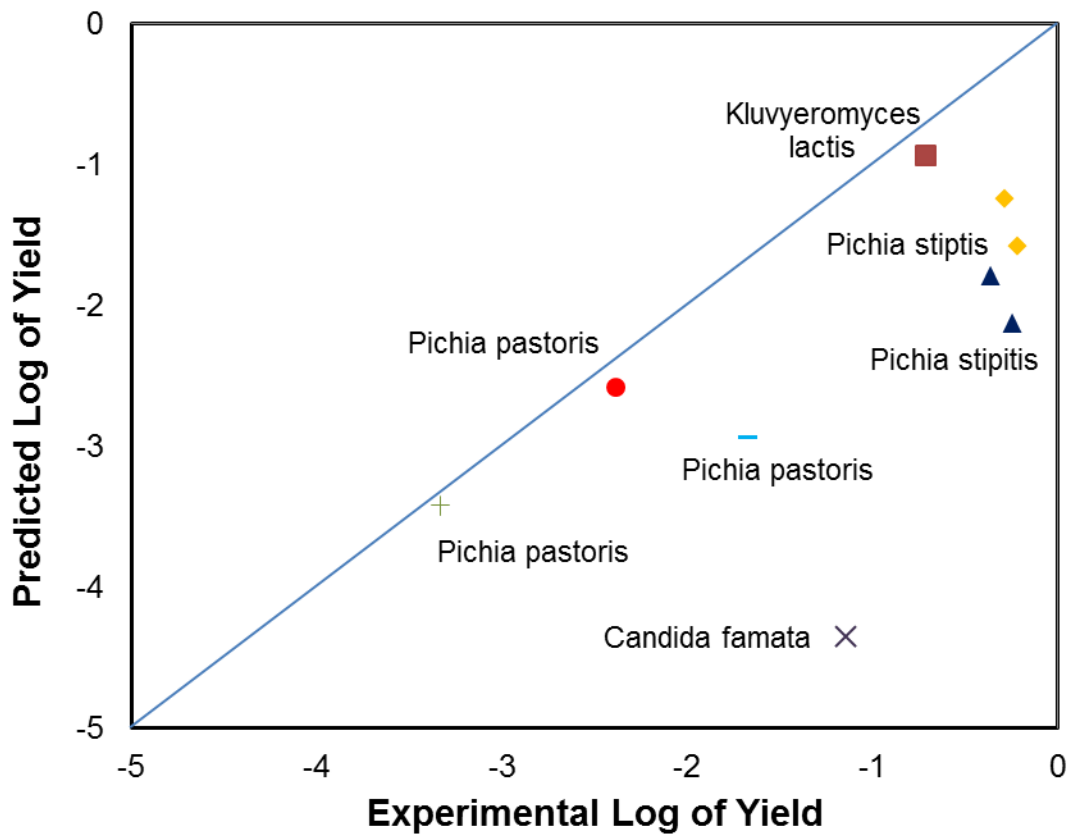


Figure 2.3: *S. cerevisiae* model prediction of biosynthesis yields for other industrial yeast species^{57-59, 68-70}. Ethanol: ■ or ◆. L-lactic acid: ▲. Lycopene: ●. Riboflavin: + or X. S-adenosyl-L-methionine: —.

Table 2.1: Ordinal variables used in the linear regression model

Ordinal variables	Category 1 (subscript C1)	Category 2 (subscript C2)	Category 3 (subscript C3)
OVE: number of modified genes or pathways	No modified genes or pathways were present.	One or two modified genes or pathways were present.	More than two modified genes or pathways were present.
KNO: number of gene knockouts in known competitive pathways	No gene knockouts were performed.	Gene knockouts were performed.	
NUT: nutrient source	Fermentation occurred in defined medium (only including trace amounts of amino acids or vitamins)	Fermentation occurred in a very rich medium.	
INT: Intermediate	Intermediate was not added	Intermediate was added	
CUL: cultivation mode	Fermentation occurred in a shaking flask.	Fermentation occurred in a batch, fed-batch, or continuous feed bioreactor.	
OXY: oxygen conditions	Fermentation occurred in aerobic conditions.	Fermentation occurred under oxygen-limited conditions (anaerobic or micro-aerobic).	

Note: the input of ordinal variables was specified using a binary system, 1 and 0. When a category (e.g., overexpression Category 2) was applied, the value 1 was assigned to OVE_{C2} . Otherwise, the value 0 was assigned.

Table 2.2: Dataset used for linear regression

Reference	Product	Yield	Primary Step	Second Step	OVE_C2	OVE_C3	KNO_C2	NUT_C2	INT_C2	CUL_C2	OXY_C2
71	(E,E,E)-Geranylgeraniol	0.00025	10	10	1	0	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.014	10	10	0	1	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.047	10	10	0	1	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.018	10	10	0	1	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.031	10	10	0	1	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.058	10	10	0	1	0	1	0	0	0
71	(E,E,E)-Geranylgeraniol	0.14	10	10	0	1	0	1	0	1	0
72	1,2-Propanediol	0.014	4	3	1	0	0	1	0	0	0
73	1,2-Propanediol	0.010	4	3	1	0	0	1	0	1	0
73	1,2-Propanediol	0.026	4	3	1	0	0	1	0	1	0
74	5-epi-aristolochene	0.010	10	9	1	0	1	1	0	0	0
74	5-epi-aristolochene	0.0090	10	9	1	0	1	1	0	0	0
75	Acetate	0.13	9	2	0	0	1	0	0	1	0
76	Acetate	0.015	9	2	0	0	1	0	0	1	0
77	Acetate	0.26	9	2	0	1	0	0	0	0	1
78	Amorphadiene	0.00049	12	9	1	0	0	0	0	0	0
78	Amorphadiene	0.0020	12	9	1	0	0	0	0	0	0
78	Amorphadiene	0.0040	12	9	1	0	1	0	0	0	0
78	Amorphadiene	0.011	12	9	1	0	1	0	0	0	0
78	Amorphadiene	0.016	12	9	0	1	1	0	0	0	0
78	Amorphadiene	0.016	12	9	0	1	1	0	0	0	0
79	Amorphadiene	0.0080	12	9	1	0	1	0	0	0	0
79	Amorphadiene	0.0090	12	9	0	1	1	0	0	0	0
79	Amorphadiene	0.011	12	9	0	1	1	0	0	0	0

79	Amorphadiene	0.013	12	9	0	1	1	0	0	0	0
78	Artemisinic acid	0.0030	12	10	0	1	1	0	0	0	0
78	Artemisinic acid	0.011	12	10	0	1	1	0	0	1	0
80	Cyanophycin	0.12	10(0)	2(1)	1	0	1	1	1	0	0
80	Cyanophycin	0.10	10(0)	2(1)	1	0	1	1	1	0	0
80	Cyanophycin	0.15	10(0)	2(1)	1	0	1	1	1	0	0
81	Dihydroxyacetone	0.0040	4	3	1	0	0	0	0	0	0
81	Dihydroxyacetone	0.034	4	3	1	0	1	0	0	0	0
82	D-Lactic acid	0.61	9	1	1	0	1	1	0	0	1
83	Dolichol	0.00010	10	11	0	0	0	1	0	0	0
83	Dolichol	0.00018	10	11	1	0	0	0	0	0	0
83	Ergosterol	0.00015	10	21	0	0	0	1	0	0	0
83	Ergosterol	0.00020	10	21	1	0	0	0	0	0	0
50	Ethanol	0.55	9	2	1	0	0	0	0	1	1
50	Ethanol	0.47	8	2	0	1	0	0	0	1	1
84	Ethanol	0.080	8	2	0	1	0	0	0	1	0
84	Ethanol	0.12	8	2	0	1	0	0	0	1	1
84	Ethanol	0.15	8	2	0	1	0	0	0	1	1
85	Ethanol	0.53	9	2	1	0	0	1	0	0	0
85	Ethanol	0.20	9	2	1	0	0	1	0	0	0
85	Ethanol	0.47	9	2	1	0	0	1	0	0	0
85	Ethanol	0.42	9	2	0	0	0	1	0	0	0
85	Ethanol	0.36	9	2	0	0	0	1	0	0	0
76	Ethanol	0.44	9	2	0	0	1	0	0	1	0
76	Ethanol	0.32	8	2	0	0	1	0	0	1	1
86	Ethanol	0.52	9	2	1	0	0	0	0	0	0
77	Ethanol	0.55	9	2	0	1	0	0	0	0	1
77	Ethanol	0.39	9	2	0	1	0	0	0	0	1
77	Ethanol	0.51	9	2	0	1	1	0	0	0	1
87	Ethylene*	0.00069	13	10	1	0	0	0	0	1	0

74	Farnesol	0.036	10	9	0	0	1	1	0	0	0
88	Flavanones	0.030	10(0)	14(3)	0	1	0	0	1	0	0
88	Flavanones	0.053	10(0)	14(3)	0	1	0	0	1	0	0
89	Formate	0.00024	6	7	0	0	1	0	0	0	0
89	Formate	0.00030	6	7	0	0	1	0	0	0	0
90	Geraniol	0.00011	10	8	1	0	0	0	0	0	0
90	Geraniol	0.00019	10	8	1	0	1	0	0	0	0
90	Geraniol	0.00019	10	8	1	0	1	0	0	0	0
91	Glycerol	0.12	4	2	1	0	0	0	0	1	0
91	Glycerol	0.12	4	2	1	0	0	0	0	1	0
75	Glycerol	0.41	4	2	1	0	1	0	0	1	0
75	Glycerol	0.45	4	2	1	0	1	0	0	1	0
75	Glycerol	0.45	4	2	1	0	1	0	0	0	0
92	Glycerol	0.49	4	2	0	0	1	1	0	1	0
92	Glycerol	0.41	4	2	0	0	1	1	0	1	0
76	Glycerol	0.050	4	2	0	0	1	0	0	1	0
76	Glycerol	0.037	2	4	0	0	1	0	0	1	1
93	Glycerol	0.45	4	2	0	0	1	0	0	1	0
93	Glycerol	0.54	4	2	1	0	1	0	0	1	0
60	Glycerol 3-phosphate	0.0010	4	1	1	0	1	0	0	0	1
94	Hydrocortisone	0.0020	10(0)	19(2)	1	0	0	1	1	0	0
94	Hydrocortisone	0.0020	10(0)	19(2)	1	0	1	1	1	0	0
94	Hydrocortisone	0.021	10(0)	19(2)	1	0	1	1	1	0	0
94	Hydrocortisone	0.026	10(0)	19(2)	1	0	1	1	1	0	0
95	Lactate	0.44	9	1	1	0	0	1	0	1	0
96	Lactate	0.21	9	1	1	0	0	0	0	1	0
97	L-Ascorbic acid	0.14	2(0)	8(2)	1	0	0	0	1	0	0
97	L-Ascorbic acid	0.066	2(0)	8(2)	1	0	0	0	1	0	0
90	Linalool	8.8×10^{-5}	10	8	1	0	0	0	0	0	0
90	Linalool	2.3×10^{-5}	10	8	1	0	1	0	0	0	0

98	L-Lactic Acid	0.65	9	1	1	0	1	1	0	0	0
62	Malate	0.28	11	0	0	1	0	0	0	0	0
79	Mevalonate	0.022	12	3	0	1	1	0	0	0	0
79	Mevalonate	0.022	12	3	0	1	1	0	0	0	0
99	Naringenin	0.0070	8(0)	15(3)	0	1	0	0	1	0	0
99	Naringenin	0.0020	8(0)	15(5)	0	1	0	0	1	0	0
44	Naringenin	0.00058	10	14	0	1	0	1	0	0	0
23	n-Butanol	0.00020	12	6	0	1	0	0	0	0	1
99	p-Coumaric Acid	0.033	8(0)	12(2)	0	1	0	0	1	0	0
100	p-Hydroxycinnamic acid	0.00020	8	12	1	0	0	0	0	0	0
100	p-Hydroxycinnamic acid	0.20	8(0)	12(2)	1	0	0	0	1	0	0
44	Pinocembrin	6.6x10 ⁻⁵	10	14	0	1	0	1	0	0	0
101	Poly[(R)-3-hydroxybutyrate]	0.00056	10	3	1	0	0	0	0	0	0
101	Poly[(R)-3-hydroxybutyrate]	0.003	10	3	1	0	0	0	0	0	0
101	Poly[(R)-3-hydroxybutyrate]	0.012	10	3	0	1	0	0	0	0	0
101	Poly[(R)-3-hydroxybutyrate]	0.00047	10	3	1	0	0	0	0	1	0
101	Poly[(R)-3-hydroxybutyrate]	0.0090	10	3	1	0	0	0	0	1	0
101	Poly[(R)-3-hydroxybutyrate]	0.018	10	3	0	1	0	0	0	1	0
101	Poly[(R)-3-hydroxybutyrate]	0.0010	10	3	0	1	0	0	0	1	1
101	Poly[(R)-3-hydroxybutyrate]	0.017	10	3	0	1	0	1	0	0	0
74	Premnaspirodiene	0.011	10	9	1	0	1	1	0	0	0
74	Premnaspirodiene	0.0090	10	9	1	0	1	1	0	0	0
102	Pyruvate	0.55	9	0	0	0	1	0	0	1	0
76	Pyruvate	0.0050	9	0	0	0	1	0	0	1	0
103	Reticuline	0.051	8(0)	16(3)	0	1	0	0	1	0	0
104	Ribitol	0.0020	5	2	0	0	1	0	0	0	0
104	Ribitol	0.027	5	2	1	0	1	0	0	0	0
104	Ribitol	0.017	5	2	1	0	1	0	0	0	0
104	Ribitol	0.021	5	2	1	0	1	0	0	0	0
71	Squalene	0.042	10	9	1	0	0	1	0	0	0

105	Taxadiene	7.7×10^{-5}	12	8	0	1	0	1	0	0	0
106	Vanillin	0.0030	3	6	0	1	1	1	0	0	0
104	Xylitol	0.0070	5	2	1	0	1	0	0	0	0
104	Xylitol	0.014	5	2	1	0	1	0	0	0	0
104	Xylitol	0.014	5	2	1	0	1	0	0	0	0
77	Xylitol	0.27	5	2	0	1	0	0	0	0	1
77	Xylitol	0.29	5	2	0	1	1	0	0	0	1
107	β -carotene	4.5×10^{-7}	10	14	1	0	0	0	0	0	0
107	β -carotene	2.9×10^{-6}	10	14	0	1	0	0	0	0	0
107	β -carotene	0.00011	10	14	0	1	0	0	0	0	0
107	β -carotene	0.00036	10	14	0	1	0	0	0	0	0
107	β -carotene	0.0010	10	14	0	1	0	0	0	0	0

Note: Some papers show that product biosynthesis can be enhanced by supplementing additional precursors. In the parenthesis, we have listed the number of enzyme steps from the added intermediates to final products.

* Steps for ethylene were counted based on the arginine route.

Table 2.3: Regression coefficients and P-values for *S. Cerevisiae* model

Variable	Model 1			Model 2			Model 3		
	With primary steps			Without primary steps			Ethanol as a primary metabolite		
	Coefficient	P-value	Std. Error	Coefficient	P-value	Std. Error	Coefficient	P-value	Std. Error
Intercept	-1.53	0	0.42	-1.60	0	0.34	-1.73	0	0.41
Primary step	-0.01	0.76	0.04	-	-	-	0.003	0.93	0.03
Secondary step	-0.19	0	0.02	-0.19	0	0.02	-0.19	0	0.02
OVE_{c2}	0.007	0.98	0.26	0.0003	0.99	0.25	0.05	0.84	0.24
OVE_{c3}	0.52	0.07	0.29	0.50	0.079	0.28	0.56	0.05	0.28
KNO_{c2}	0.31	0.08	0.18	0.31	0.078	0.18	0.37	0.03	0.17
NUT_{c2}	0.73	0	0.18	0.73	0	0.18	0.71	0	0.17
INT_{c2}	0.77	0.02	0.31	0.82	0.001	0.25	0.86	0.004	0.29
CUL_{c2}	0.51	0.02	0.22	0.51	0.02	0.21	0.51	0.02	0.21
OXY_{c2}	0.27	0.32	0.27	0.28	0.31	0.27	0.12	0.65	0.27
Multiple R²		0.55			0.55			0.58	

Chapter 3: The use of ^{13}C -based analysis to elucidate the intrinsic biosynthesis yields

This chapter has been submitted for peer-review and the author would like to thank all the co-authors for their contributions.

Abstract

Microbial platforms have widely been used for the synthesis of diverse value-added chemicals. Rational metabolic engineering and optimal fermentations can improve microbial cell factory's yields from renewable feedstock. However, very few studies have rigorously investigated the intrinsic product yields from engineered microbial platforms under complex cultivation conditions. In this paper, we discuss the use of ^{13}C -based metabolite analysis for assessment of product yields in four different cases. First, in the rich medium fermentation, undefined nutrients (such as yeast extract) may also contribute to the synthesis of final product. Second, product synthesis may be dependent on co-metabolism of multiple-feedstock. Third, multiple pathways may be employed by microbes for product synthesis. Fourth, the loss of ATP/NADH due to cell maintenance and low P/O ratio (Phosphate/Oxygen Ratio) reduces product yields, while ^{13}C -metabolic flux analysis (^{13}C -MFA) can assess the influence of suboptimal energy metabolism on microbial productivity. Since product yield is a major determining factor in biotechnology commercialization, we foresee that ^{13}C -isotopic labeling experiments, even without performing extensive ^{13}C -flux calculations, can play valuable roles in the development of microbial cell factories.

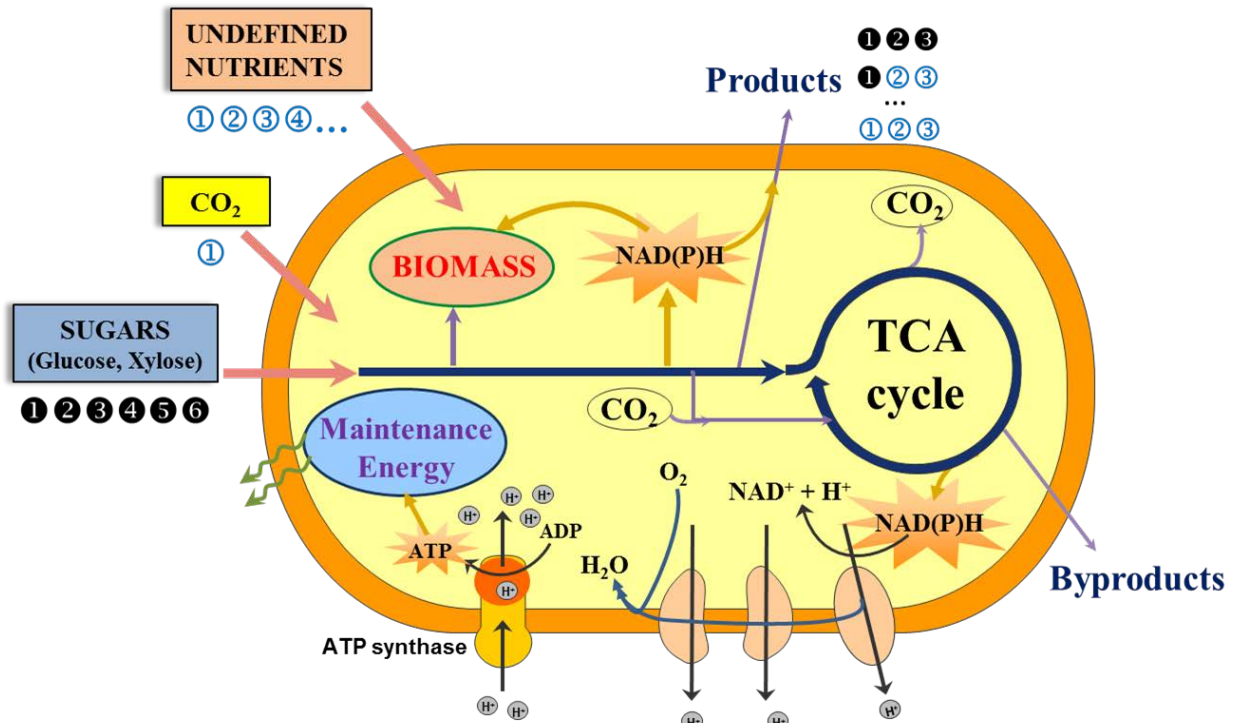
Keywords: cell maintenance, co-metabolism, metabolic flux analysis, P/O ratio, yeast extract

3.1 Introduction

Recent advances in metabolic engineering has enabled us to engineer microbial cell factories for the efficient synthesis of diverse products, including bulk chemicals, drugs and fuels¹⁰⁸. For example, advanced biofuels produced by engineered microorganisms with properties similar to that of petroleum-based fuels, are being reported extensively^{25, 109-112}. The emergence of systems biology and synthetic biology has greatly increased the potential of microbial cell factories towards the production of value-added chemicals¹¹³. This has also improved product's yield, titer, and rate so that microbial cell factories can be moved from lab scale to industrial fermentations^{114, 115}. The product yield is a key indicator in achieving an economical bio-production of the bulk and commodity chemicals¹¹⁶. But, estimation of product yield may be difficult if fermentations use either rich-mediums or multiple feedstock (Figure 3.1). Moreover, new pathways/enzymes are often employed to improve microbial productivity and their relative contribution to product yield remains unknown^{24, 109, 117, 118}. Thereby, a proper technique for the quantification of intrinsic yields from the engineered pathways is needed if multiple biosynthesis routes are used by microbial hosts.

¹³C-tracing experiments can rigorously determine the *in vivo* carbon fluxes from specific substrates to final products. Feeding microbial cultures with ¹³C-labeled substrates results in unique isotopic patterns amongst the cell metabolites (isotopic fingerprints)¹¹⁹, which can provide functional characterization of metabolic pathways¹²⁰. Integration of this isotopomer data with metabolic modeling (i.e., ¹³C-MFA) can be used to predict cellular metabolic fluxes. The metabolic fluxes not only reveals metabolic responses to product synthesis and growth conditions^{121, 122}, but can also reveal the rigid metabolic nodes for rational pathway engineering

1 ¹²³. Therefore, ¹³C-based analysis (i.e., pathway tracing and ¹³C-MFA) are widely used in the
 2 field of biotechnology ¹²⁴⁻¹²⁷. Besides these common applications, this paper demonstrates the
 3 additional utility of simple ¹³C-experiments or the more complicated ¹³C-MFA in determining
 4 product yields from microbial cell factories.



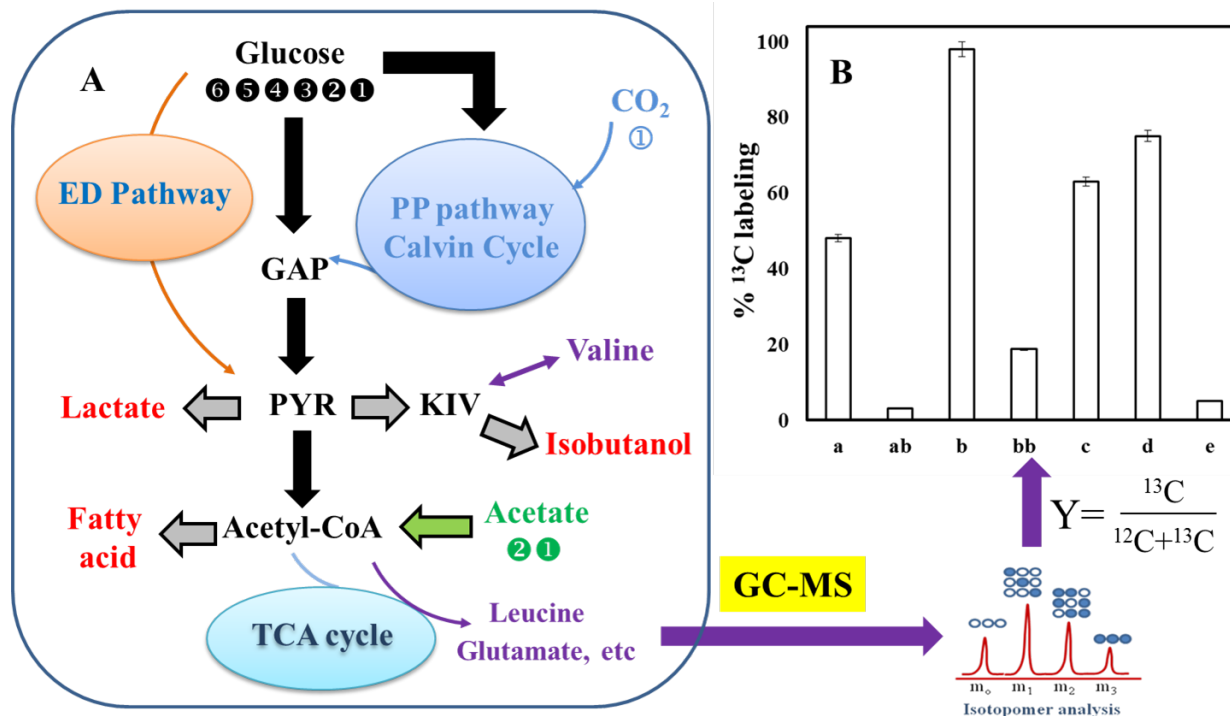
5 **Figure 3.1:** Schematic description of microbial metabolism. Many microbes have the ability to
 6 co-metabolize diverse feedstock. Dark circles indicate labelled carbon. The enrichment of
 7 labeling in the product acts as an indicator for the relative uptake flux of sugars.
 8
 9

10 3.2 Product yield using rich medium

11
 12 Engineered microbes may have many metabolic burdens that can inhibit both biomass
 13 growth and product synthesis. To promote their productivity, rich media are commonly used in
 14 fermentations as they provide diverse nutrients for cell growth and stabilize the production
 15 performance of the microbe ^{114, 115}. Thereby, rich mediums include both primary carbon
 16 substrates (e.g., sugars) and large amounts of nutrients (such as yeast extract). Multiple studies

1 have revealed that supplementing culture medium with yeast extract or terrific broth to
2 engineered microbes significantly improves their final biosynthesis yields^{63, 128}. Since nutrient
3 supplements can provide undefined building blocks for both biomass and product synthesis, it is
4 difficult to precisely calculate the actual product yield from the rich-medium fermentation. To
5 overcome this problem, ¹³C-analysis can be employed to gain insights into the contributions of
6 nutrients to product biosynthesis.

7 For example, two *E. coli* strains engineered for isobutanol production (i.e., a low
8 performance strain with an Ehrlich pathway¹²⁹ and a high performance JCL260 strain with
9 overexpression of both the keto-acid pathway and the Ehrlich pathway¹³⁰) display an increase in
10 isobutanol titer with the inclusion of yeast extract in their culture medium. Using fully labeled
11 glucose and non-labeled yeast extract as carbon sources, ¹³C-experiments revealed that the low-
12 performance strain derived ~50% of isobutanol carbons from yeast extract (Figure 3.2), while the
13 JCL260 strain synthesized isobutanol solely from ¹³C-glucose and used yeast extract mainly for
14 biomass growth¹²⁹. This observation from ¹³C-analysis indicates that overexpression of keto-
15 acid pathway can resolve the isobutanol synthesis bottleneck and effectively pull the carbon from
16 glucose to product. In another work, an *E.coli* strain was engineered for conversion of acetate
17 into free fatty acids via the overexpression of both acetyl-coA synthetase and the fatty acid
18 pathways. In the acetate-based fermentation, yeast extract significantly promoted fatty acid
19 productivity, resulting in 1 g/L fatty acids from ~10g/L acetate¹³¹. ¹³C-analysis of the culture
20 with fully labeled acetate and yeast extract has shown that ~63% carbons in the free fatty acids
21 were synthesized from ¹³C-acetate (Figure 3.2). Thereby, the actual microbial yield from a
22 primary substrate in a rich medium could be correctly estimated based on isotopomer analysis.



1
 2 **Figure 3.2:** (A) Biosynthesis yield analyzed by feeding cells with ^{13}C -substrates (such as fully
 3 labeled glucose and acetate). (B) Relative product yields from a primary substrate (a –
 4 Isobutanol from glucose in a low performance strain; ab – valine from glucose in a low
 5 performance strain; b – Isobutanol from glucose in JCL260; bb – valine from glucose in JCL260)
 6 129 ; c – Free fatty acids from acetate in an *E.coli* strain 131 ; d - biomass from glucose in wild type
 7 *Synechocystis* 6803 132 ; e - D-lactate from acetate in engineered *Synechocystis* 6803 133 . Relative
 8 yield is calculated based on ^{13}C concentrations in the final product. Abbreviations: GAP,
 9 Glyceraldehyde -3- phosphate; PYR, pyruvate; KIV, ketoisovalerate.

10

11 3.3 Product yield during co-metabolism of multiple carbon substrates

12

13 Algal species are able to utilize both CO_2 and organic carbon substrates. Such
 14 mixotrophic metabolism can alleviate the dependence of algal hosts on light and CO_2 limitations,
 15 and thus enable them to achieve high biomass growth rate and product titer 134 . ^{13}C -metabolite
 16 analysis has been used to track their photomixotrophic metabolisms in different scenarios. For
 17 example, *Synechocystis* sp. PCC 6803 (blue-green algae) is a model cyanobacterium which can
 18 be engineered to produce diverse products and has capability to perform photomixotrophic

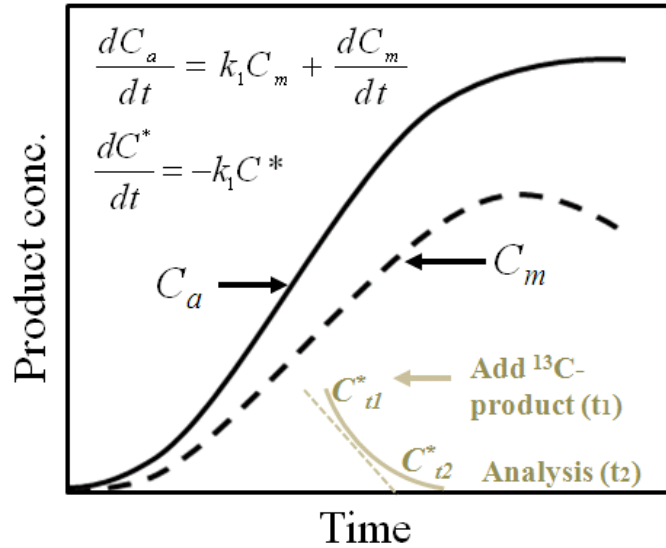
1 metabolism ¹³⁵. ¹³C-MFA has shown that CO₂ contributes to 25% of *Synechocystis* biomass yield
2 during its mixotrophic growth with ¹³C-glucose and ¹²CO₂ ¹³². Besides, ¹³C-metabolite analysis
3 has also been used to track the synthesis of D-lactate in an engineered *Synechocystis* 6803 ¹³³. In
4 that study, the lactate production was found to be increased substantially during the co-
5 metabolism of both CO₂ and acetate. Experiments with fully labeled acetate and ¹²CO₂
6 discovered that nearly all of lactate molecules were non-labeled and that only the acetyl-CoA-
7 derived proteinogenic amino acids (leucine, glutamate and glutamine) were ¹³C-labeled. This
8 result suggested that acetate was involved only in biomass growth, while the yield of D-lactate
9 was completely derived from CO₂. ¹³C-results also further indicated that acetate may inhibit the
10 pyruvate decarboxylation reaction and thus redirect flux to lactate. The above study shows the
11 value of ¹³C-analysis to improve our understanding of pathway regulations for product synthesis.
12 Since many microbial platforms (including both algal species and heterotrophs) may co-
13 metabolize multiple carbon substrates simultaneously, isotopomer feeding can reveal the
14 contributions of each substrate to key metabolite pools, and thus predict the potential bottlenecks
15 in biomass or product formations.

16 **3.4 Accurate laboratory analysis of product concentrations**

17

18 Direct measurement of product concentrations in the culture can obtain deceptive results.
19 There are a few cases that cause product measurement errors. First, loss of volatile products
20 (such as alcohols) during fermentation may reduce product titers. Second, product may be
21 degraded or consumed by contaminated microbes during fermentation process. For example,
22 photochemical degradation of isobutanol synthesis from cyanobacteria was reported ¹³⁶. Third,
23 aerobic fermentation in shake flasks may have significant water vaporization during long-term

1 incubation because the cultivation volume is relatively small (<50mL working volume) and thus
 2 the product concentration may be condensed (e.g., 10~20% water loss was normally observed
 3 after three-day shaking flask cultivations at 37 °C). In all these circumstances, the final product
 4 yield could be very different from the intrinsic microbial product yield. To obtain the true
 5 productivity, *in situ* product recovery is a common method to reduce product loss. For example,
 6 volatile alcohol products can be trapped in organic solvents during microbial fermentation (e.g.,
 7 gas stripping)^{130, 136, 137}. Alternatively, kinetic model can be used to obtain intrinsic product yields
 8 via parameter estimation based on complete time-course fermentation data and statistical
 9 analysis (to avoid local solutions)¹²⁹. Thirdly, ¹³C-experiments can also resolve artifacts during
 10 measurement of intrinsic product yields. By addition of small amount of ¹³C-product in the
 11 culture as an internal standard, we can directly measure the change of ¹³C-product during
 12 fermentation (Figure 3.3). Then, ¹³C-data from time points can be used to correct the artifact of
 13 yield coefficients: $\frac{dC^*}{dt} = -k_1 C^*$, where C^* represents the ¹³C concentration of the product and
 14 k_1 is assumed the first order constant for product loss. Thereby, $k_1 = -\ln\left(\frac{C_{t_2}^*}{C_{t_1}^*}\right) / (t_2 - t_1)$. By
 15 analyzing the change in ¹³C concentration at two different time points (t_1 and t_2), k_1 can be
 16 calculated. The product loss term can be added to the normally measured product curves for
 17 correction of intrinsic product curves.



1

2 **Figure 3.3:** Schematic showing the dynamics of product concentration. C_a is the actual
 3 concentration and C_m is the measured product concentration. C^* indicates the concentration of
 4 the ^{13}C product added as an internal standard.

5

6 **3.5 Assessment of maximum product yield**

7

8 Theoretical product yield is generally calculated based on the stoichiometry of product
 9 synthesis from a carbon substrate (without accounting both biomass growth and waste secretion).

10 However, microbial energy metabolism may also be affecting product yield which is seldom

11 accounted. The synthesis of high-energy chemicals often requires large amounts of ATPs, while

12 cell maintenance (used for regeneration of degraded macromolecules, futile cycles, and ATP

13 leaks) also competes for the same ATPs ¹³⁸. Oxidative phosphorylation of NADH is a major

14 source for ATP generation (theoretical P/O ratio: 1 NADH \rightarrow 3 ATPs) ¹³⁹. However, respiration

15 efficiency in engineered strains could be poor (e.g., the P/O ratio = 1.3 during riboflavin

16 fermentation ¹⁴⁰) due to metabolic stresses ¹⁴¹. Thereby, a cell may consume extra substrates to

17 compensate for the ATP demand. To illustrate the effect of cell maintenance on product yield ¹⁴²,

18 a flux balance model was built to show free fatty acid production as a function of ATP

1 maintenance and P/O ratios (Figure 3.4) ¹⁴³. This model employs eight reactions (Table 3.1) and
2 the fluxes were resolved by the function below:

$$3 \quad \left[\begin{array}{l} \max v(2) \\ \text{such that } A \cdot v = b \text{ and } lb \leq v \leq ub \end{array} \right],$$

4 where the objective function is to maximize $v(2)$ (i.e., the relative flux of fatty acid). A is the
5 reaction stoichiometry. lb and ub are upper and lower bound for each reaction flux, $v(i)$. Fig.
6 3.4a shows the relationship between maximum yield, P/O ratio and ATP maintenance without
7 biomass growth ($v(8)=0$). A Higher P/O ratio makes the microbial system less sensitive to the
8 increased demand for ATP. When the ATP maintenance is low and the P/O ratio is close to 3, the
9 fatty acid yield can reach the theoretical value of 0.36g fatty acid/g glucose. In this case,
10 reduction of carbon loss via knock-out of competitive pathways will be effective to achieve the
11 theoretical yield (Figure 3.4A). If ATP consumption for maintenance increases, cells need to
12 “burn” extra carbon substrates for energy generation so that fatty acid yields drop significantly.
13 In this case, the biosynthesis optimization needs to reduce the loss of ATP/NADH. For example,
14 in a study of the engineered *E.coli* metabolism responding to fatty acid overproduction ¹⁴³, ¹³C-
15 MFA (via extensive flux calculation) found that the total ATP/NAD(P)H generation was much
16 higher than their consumption for biomass growth and fatty acid synthesis. Such difference was
17 attributed to the low P/O ratio and high cell maintenance during fatty acid overproduction.
18 Therefore, the engineered strain had a fatty acid yield of only 0.17g fatty acid/g glucose even
19 after extensive pathway engineering, (Figure 3.4B). The suboptimal energy metabolism in the
20 engineered strain was likely due to the various physiological stresses experienced by the cell
21 during fatty acid overproduction (e.g., change cell membrane integrity and compositions ¹⁴⁴).
22 Since metabolic stresses are commonly experienced by microbial hosts, ¹³C-MFA can provide a

1 diagnostic analysis of cell maintenance and offer insights into the metabolic potential for
 2 improving biosynthesis¹⁴⁵.

3

4 **Table 3.1:** Simplified biochemical reactions considered in the model

5

Flux, v	Reactions	Note
v(1)	Glucose \rightarrow 2AceCoA + 2ATP + 4NADH	Glycolysis
v(2)	AceCoA + 1.75NADPH + 0.875ATP \rightarrow 0.125 C16:0 fatty acid	Fatty acid synthesis
v(3)	AceCoA \rightarrow 2NADH + NADPH + ATP + FADH2	TCA cycle
v(4)	NADH \rightarrow NADPH	Transhydrogenation
v(5)	NADH \rightarrow P/O ATP	Oxidative phosphorylation
v(6)	FADH2 \rightarrow 0.67(P/O)ATP	Oxidative phosphorylation
v(7)	ATP \rightarrow ATP_ext	ATP maintenance
v(8)	6.6Glucose + 37.6ATP + 9.5NADPH + 2.5AceCoA = 39.7Biomass + 3.1NADH	Biomass formation

6 Note: glucose consumption for both biomass growth and product synthesis is normalized to 100.

7 The optimization was performed by a linear optimizer 'linprog' in MATLAB. The final yield (g
 8 fatty acid/g glucose) is calculated as follows: $Y=(v(2)/8\cdot 256)/(100\cdot 180)$ g C16:0 fatty acid/g
 9 glucose.

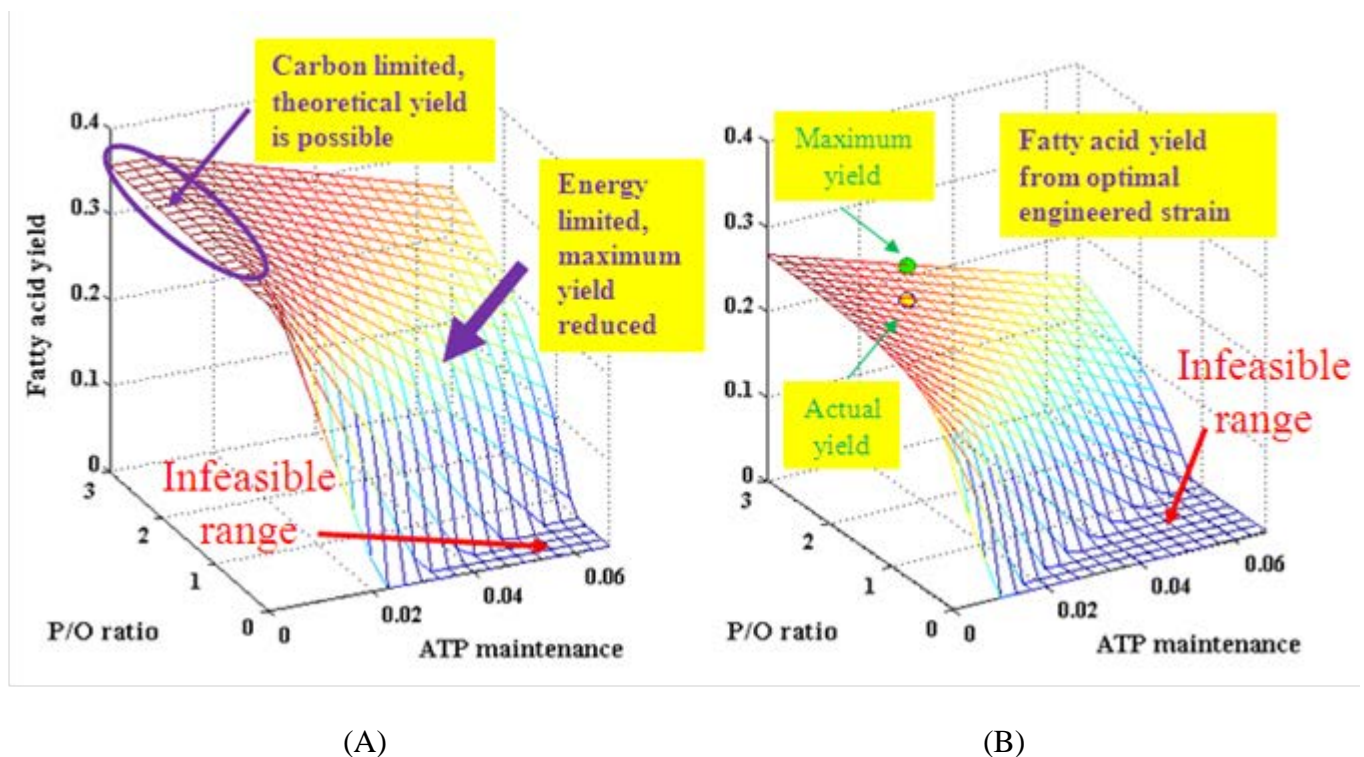


Figure 3.4: (A) Theoretical Yield as a function of P/O ratio and ATP maintenance without biomass growth. (B) Theoretical Yield as a function of P/O ratio and ATP maintenance at growth rate $v(8)=3.6$. The units of yield and ATP maintenance are ‘g C16:0 fatty acid/g glucose’ and ‘mol ATP /g glucose’ respectively. The infeasible range in the surface plot indicates that, energy cannot be balanced for fatty acid or biomass production in that region, resulting in zero yield¹⁴³.

3.6 Product yield from unconventional engineered pathway

¹³C-analysis can be used to decipher the yield of products through multiple biosynthesis routes. For example, the acetogenic bacterium *Clostridium carboxidivorans* uses syngas (H₂, CO and CO₂) to generate various chemicals (e.g., acetate, ethanol, butanol, and butyrate)¹⁴⁶. It contains several routes for CO₂ fixation, which includes the Wood-Ljungdahl pathway and the anaplerotic or the pyruvate synthase reactions. ¹³C-experiments can be used to identify the relative contribution of the different CO₂ fixation pathways towards product synthesis. As a demonstration, cultivation of *Clostridium* with labeled ¹³CO₂ and ¹²CO has been shown in Fig.

3.5. Here, ^{13}C -analysis of the labeling patterns in either alanine or pyruvate can reveal the relative contributions of the different CO_2 assimilation routes to biomass and product synthesis.

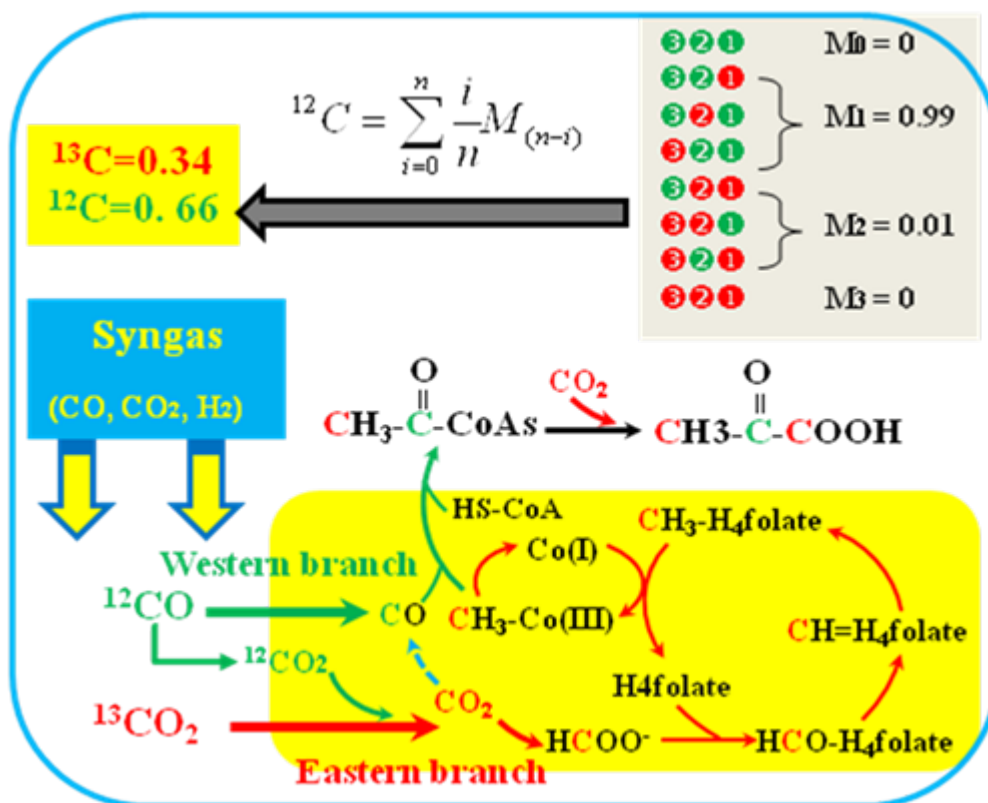


Figure 3.5: ^{13}C analysis to study the carbon assimilation during syngas fermentation ($^{13}\text{CO}_2$, ^{12}CO and H_2). Analysis of metabolite labeling patterns can determine CO_2 and CO utilization for pyruvate production. The isotopomer data of pyruvate were used as a demonstration of ^{13}C applications for product yield calculations.

“Rule of Thumb” indicates that 20%~30% yield reduction happens per engineered enzymatic reaction step^{63, 128}. Thereby, novel pathways are constantly being explored and engineered into microbial hosts to create a short-cut route from the feedstock to the final product. If new pathways are engineered into microbes, it could be unclear how much the engineered pathways are used by the microbe in parallel with its original pathway¹⁴⁷. In the following example, we demonstrate that ^{13}C -experiments can determine the relative fluxes through multiple

pathways based on measurements of the product labeling. Specifically, butanol could be

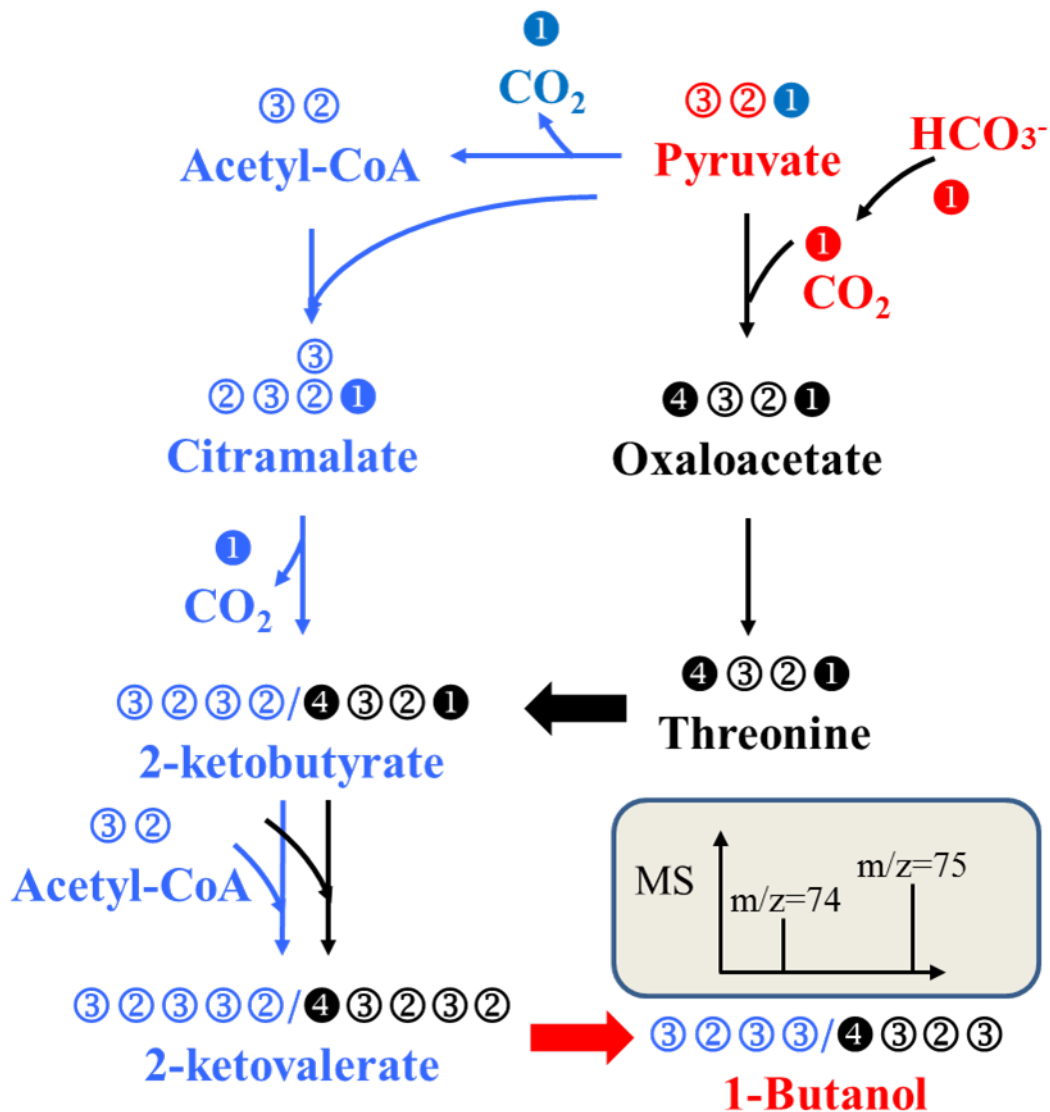


Figure 3.6: Threonine and citramalate pathway for the synthesis of 1-butanol. The carbon rearrangement network shows the labeling of 1-butanol for both the pathways, when fed with 1- ^{13}C pyruvate and ^{13}C bicarbonate.

produced simultaneously from a threonine pathway and a citramalate pathway (a short-cut keto acid-mediated pathway) in *E.coli*¹⁴⁸. If 1st position ^{13}C -pyruvate and ^{13}C -bicarbonate are fed to the butanol producing strain, labeling patterns in butanol can quantify fluxes through the two

biosynthesis pathways (Fig. 3.6). In another example, James Liao Lab introduced a non-oxidative glycolytic cycle (NOG) into *E.coli* for breaking down hexose that could lead to a 50% increase of biofuel yield¹¹⁷. This NOG pathway starts with fructose 6-phosphate and contains three metabolic cycles to generate Acetyl-CoA without carbon loss. To probe the contribution of NOG pathway to overall cell metabolism, their study has also presented a carbon rearrangement map so that ¹³C-tools can be employed. These examples illustrate that ¹³C-analysis is potentially suited to examine *in vivo* activity of these novel pathways for product synthesis.

3.7 Conclusion

Product yield is one of the main factors involved in commercialization of a technology¹⁴⁹. Microbial productivity is not only associated with the efficiency of biosynthesis enzymes, but is also intertwined with energy metabolisms and metabolic balances¹⁵⁰. Via simple ¹³C analysis, we can characterize the hosts' intrinsic production yield using different carbon sources, and determine the contributions of alternative pathways to biosynthesis. In addition, ¹³C-MFA can profile hosts' fluxomes and determine the amount of extra substrates that cell metabolism has to consume to compensate ATP losses due to cell maintenance and low P/O ratio. In the end, loss of products during fermentations (such as volatile alcohols or product degradation) introduces common measurement errors and artifacts^{130, 136, 137}. The accurate quantification of unstable metabolites can also be achieved via ¹³C-based method (i.e., using isotopomer labeled internal standards)¹⁵¹. Through this review paper, we hope that the metabolic engineering field will recognize more value of ¹³C-techniques and foresee an extended use of ¹³C-experiments during the development of microbial cell factories.

Acknowledgement

We thank Prof. Maciek Antoniewicz from University of Delaware for his advice. This work was supported in parts by the funding from National Science Foundation (MCB0954016) and the Bridge Funding from Washington University in St. Louis.

Chapter 4: Metabolic engineering of *Synechocystis* 6803 for isobutanol production

This chapter has been reproduced from the following publication:

Varman, A.M., Xiao, Y., Pakrasi, H.B. & Tang, Y.J. Metabolic Engineering of *Synechocystis* sp. Strain PCC 6803 for Isobutanol Production. *Applied and Environmental Microbiology* **79**, 908-914 (2013).

Abstract

Global warming and decreasing fossil fuel reserves have prompted great interest in the synthesis of advanced biofuels from renewable resources. In an effort to address these concerns, we have performed metabolic engineering of the cyanobacterium *Synechocystis* sp. PCC 6803 to develop a strain that can synthesize isobutanol under both autotrophic and mixotrophic conditions. With the expression of two heterologous genes from the Ehrlich Pathway, the engineered strain can accumulate 90 mg/L of isobutanol from 50 mM bicarbonate in a gas-tight shaking flask. This strain does not require any inducer (i.e., IPTG: Isopropyl β -D-1-thiogalactopyranoside) or antibiotics to maintain its isobutanol production. In the presence of glucose, isobutanol synthesis is only moderately promoted (titer = 114 mg/L). Based on isotopomer analysis, we find that compared to the wild-type strain, the mutant significantly reduced its glucose utilization and mainly employed autotrophic metabolism for biomass growth and isobutanol production. Since isobutanol is toxic to the cells and may also be degraded photochemically by hydroxyl radicals during the cultivation process, we employed *in situ* removal of the isobutanol using oleyl alcohol as a solvent trap. This resulted in a final net concentration of 298 mg/L of isobutanol under mixotrophic culture conditions.

4.1 Introduction

Global energy needs continue to increase rapidly due to industrial and development demands, furthering environmental concerns. Much of the worldwide energy consumption comes from the burning of fossil fuels, which produces about 6 gigatons of CO₂ annually¹⁵². Increasing CO₂ levels may act as a feedback loop to increase the soil emissions of other greenhouse gases such as methane and nitrous oxide, heightening global temperature¹⁵³. For energy security and environmental concerns, there is an urgent demand for the development of bioenergy. Bioethanol is the most common biofuel, but it also has low energy density and absorbs moisture. Isobutanol (IB) is a better fuel because it is less water soluble and has an energy density / octane value close to that of gasoline^{154, 155}. Amongst the next generation biofuels synthesized from pyruvate, IB possesses fewer reaction steps (5 reaction steps from pyruvate to IB) in contrast to the synthesis of 1-butanol or biodiesel. IB is less toxic to microbes²⁵ so that it may achieve higher product titer and yield^{63, 128}. For example, a maximum titer of 50.8 g/L of IB can be achieved in an engineered *E. coli*¹³⁰.

On the other hand, cyanobacteria can not only convert CO₂ into bio-products, but also can play an important role in environmental bioremediations. The photosynthetic efficiency of cyanobacteria (3~9%) is high compared to higher plants ($\leq 0.25\sim 3\%$)^{1, 13}. Furthermore, some species of cyanobacteria are amenable to genetic engineering. Table 4.1 lists the various biofuels that have been synthesized through the metabolic engineering of cyanobacteria. Autotrophic IB production in cyanobacteria was first demonstrated in *Synechococcus* 7942²⁸. Moreover, a model cyanobacterium *Synechocystis* sp. PCC 6803 is capable of growing under both photoautotrophic and mixotrophic conditions, while the presence of glucose can significantly

promote biomass and bioproduct synthesis¹⁵⁶. Thereby, we have engineered a glucose tolerant *Synechocystis* 6803 strain with two key genes *kivd* and *adhA* of the Ehrlich pathway²⁶ so that the cyanobacterial strain can convert CO₂ into IB. Through both metabolic engineering and bioprocess optimization, we have improved our strain's IB production capabilities.

4.2 Materials and methods

4.2.1 Chemicals and reagents.

Restriction enzymes, T4 DNA ligase, DNase and a Revertaid first strand cDNA synthesis kit were purchased from Fermentas or New England Biolabs. Oligonucleotides were purchased from Integrated DNA Technologies. Toluene, IB, α -ketoisovaleric acid, phenol and chloroform were purchased from Sigma-Aldrich (St. Louis, MO). KlenTaq-LA¹⁵⁷ was purchased from DNA Polymerase Technology (St. Louis, MO). TRI Reagent® was purchased from Ambion, USA. ¹³C labeled glucose was purchased from Cambridge Isotope Laboratories.

4.2.2 Culture medium and growth conditions.

A glucose tolerant wild-type strain of *Synechocystis* 6803 (WT) and the recombinant strain AV03 were grown at 30°C in liquid BG-11 medium or solid BG-11 medium at a light intensity of 50 μmol of photons $\text{m}^{-2}\text{s}^{-1}$ in ambient air. Kanamycin at a concentration of 20 $\mu\text{g}/\text{mL}$ was added to the BG-11 medium when required. Growth of the cells was monitored by measuring OD₇₃₀ of the cultures on an Agilent Cary 60 UV-Vis Spectrophotometer. Cultures for the synthesis of IB were grown in 10 mL medium (Initial OD₇₃₀ of 0.4) in 50 mL shake flasks for 4 days. The mid-log phase cultures were then closed with rubber caps to prevent the loss of IB during incubation, and the cultures were supplemented with 50 mM NaHCO₃ as an inorganic carbon source. Mixotrophic cultures of *Synechocystis* 6803 were started in a BG-11 medium

containing a known amount of glucose as an organic carbon source. *E. coli* strain DH10B was the host for all plasmids constructed in this study. *E. coli* cells were grown in falcon tubes containing Luria-Bertani (LB) medium at 37°C under continuous shaking. Ampicillin (100 µg/mL) or kanamycin (50 µg/mL) was added to the LB medium when required, for the propagation of plasmids in *E. coli*.

4.2.3 Plasmid construction and transformation of *Synechocystis* 6803.

The vector pTAC-KA containing an ampicillin resistance cassette (Amp^R) and two genes (*kivd* and *adhA* from *Lactococcus lactis*) was constructed as described¹²⁹. The pTAC-KA vector was modified using the following steps to convert it into a *Synechocystis* 6803 vector. To clone the flanking regions of a potential neutral site into pTAC-KA, a *Synechocystis* 6803 vector pSL2035 containing both the flanking regions and the kanamycin resistance cassette (Km^R) was used as a template. pSL2035 is a *Synechocystis* 6803 vector designed to integrate any foreign DNA into the genome of *Synechocystis* 6803 by replacing the *psbA1* gene and its promoter. *psbA1* is a member of *psbA* gene family and is found to be silent under most conditions^{158, 159}. pSL2035 was constructed by cloning the flanking regions for the *psbA1* gene and the Km^R into pUC118. The 5' flanking region from pSL2035 was PCR amplified along with Km^R by respective primers (Table 4.2) and cloned into the PciI and Bsu36I site of pTAC-KA, resulting in the vector pTKA2. The 3' flanking region was PCR amplified from pSL2035 by the respective primers and inserted into the AhdI site of pTKA2, disrupting the native Amp^R and henceforth creating the vector pTKA3.

Transformation was performed by using a double homologous recombination system, and the genes were integrated into the target site of the *Synechocystis* genomic DNA. Specifically, 2

mL of *Synechocystis* 6803 from a mid-log phase ($1\sim 3\times 10^8$ cells mL⁻¹) culture was centrifuged at $10,000 \times g$ for 2 min. The pellet was suspended in a fresh BG-11 medium (200 μ L) to a final cell density of $1\sim 3\times 10^9$ cells mL⁻¹. Plasmid DNA was added to a final DNA concentration of 5~10 μ g/mL¹⁹ to this dense *Synechocystis* 6803 cell culture. The mixture was then incubated under normal light conditions ($50 \mu\text{E m}^{-2} \text{s}^{-1}$) overnight. The culture was then spread onto a BG-11 agar plate containing 20 μ g/mL of kanamycin. Recombinant colonies usually appear between 7 and 10 days. Colonies were propagated on a fresh BG-11 plate containing kanamycin, and a colony PCR was performed to verify successful integration of the insert into the genomic DNA of the recombinant. The positive colonies were propagated continuously onto BG-11 plates containing kanamycin, to get a high segregation of the insert in the recombinant¹⁷. To verify the integrity of the promoter and gene sequences, the heterologous DNA integrated into the genome of the mutant AV03 was PCR amplified and sent for sequencing with the respective primers.

4.2.4 Reverse transcription PCR (RT-PCR).

Total RNA isolation of *Synechocystis* 6803 was performed using a TRI Reagent® (Ambion, USA) by following the manufacturer protocol with modifications. 1 mL of RNAwiz was prewarmed to 70°C and pipetted into the frozen cells directly. Immediately, the mixture was vortexed and incubated for 10 min at 70°C in a heater block. 0.2 mL of chloroform was added to the mixture and mixed vigorously followed by incubation at room temperature for 10 min. The aqueous and the organic phase were separated by centrifugation at $10,000 \times g$ at 4°C. The RNA containing aqueous phase was transferred into an eppendorf tube, to which, equal volumes of phenol and chloroform were added. The mixture was mixed vigorously followed by centrifugation to separate the aqueous and the organic phase. The aqueous phase was removed to a clean tube, to which 0.5 mL of diethyl pyrocarbonate (DEPC) treated water was added. RNA

in the solution was precipitated by the addition of room temperature isopropanol and centrifuged at $10,000 \times g$ at 4°C to pellet the RNA. The RNA was washed with ethanol and resuspended in a fresh $50 \mu\text{L}$ of DEPC treated water. The quantity and quality of the isolated RNA was determined using a Nanodrop ND-1000 (Thermo Scientific, USA). The RNA was incubated at room temperature with DNase to degrade any genomic DNA, if present in the RNA sample. Synthesis of cDNA was performed by utilizing a Reverse transcriptase enzyme from Fermentas along with dNTPs and random primers in a reaction buffer. The mixture was incubated at 42°C for 60 min. The synthesized cDNA was used as a template for the PCR, to detect the expression of the mRNA of interest.

4.2.5 Isobutanol quantification assay.

IB synthesized in the culture was quantified using a gas chromatograph (Hewlett Packard model 7890A, Agilent Technologies, equipped with a DB5-MS column, J&W Scientific) and a mass spectrometer (5975C, Agilent Technologies). IB extraction was done using a modified procedure¹¹⁸. Samples of the cyanobacterial culture ($400 \mu\text{L}$) were collected and centrifuged at $10000 \times g$ for 5 min. IB was extracted from the supernatant by vortexing for 1 min with $400 \mu\text{L}$ of toluene, and methanol was used as the internal standard. A $1 \mu\text{L}$ sample of the organic layer was injected into the gas chromatograph (GC) with helium as the carrier gas. The GC oven was held at 70°C for 2 min and then raised to 200°C with a temperature ramp of $30^{\circ}\text{C min}^{-1}$, and the post run was set at 300°C for 6 min. The range of the mass spectrometer (MS) scan mode was set between m/z of 20 and 200. The concentration of IB present in the culture was determined based on a calibration curve prepared with known concentrations of IB ranging from 25 mg/L to 400 mg/L .

4.2.6 ¹³C-experiment to detect carbon contribution of glucose.

The ¹³C-abundance of some important metabolites was measured for both the wild-type and the mutant strain AV03, to estimate the carbon contribution of both glucose (fully labeled by ¹³C) and nonlabeled bicarbonate for biomass and IB synthesis. Mixotrophic cultures of both the wild-type *Synechocystis* 6803 and the mutant AV03 were grown in BG-11 medium (with 50mM nonlabeled NaHCO₃), which contained 0.5% glucose (U-¹³C, Cambridge Isotope Laboratories, MA). Cultures were collected on day 3, 6 and 9, and proteinogenic amino acids were hydrolyzed and then derivatized with TBDMS (N-(tert-butyldimethylsilyl)-N-methyl-trifluoroacetamide, Sigma-Aldrich). The derivatized amino acids were analyzed for their mass isotopomer abundance by GC-MS, as described before ^{160, 161}. The m/z ion [M-57]⁺, which corresponds to the entire amino acid, was used to calculate the ¹³C abundance in amino acids [m₀ m₁... m_n]. The fraction of carbon (F_A) derived from fully labeled glucose for each amino acid was estimated

based on the following equation:
$$F_A = \frac{\sum_{i=1}^n i \times m_i}{n} \quad (4.1)$$

where i is the number of labeled carbons, m_i is the mass fraction for different isotopomers of the corresponding amino acid and n represents the total number of carbons in the corresponding amino acid. The m/z of [M-15]⁺ was used only for leucine and isoleucine, since their [M-57]⁺ overlaps with other mass peaks ¹⁶². IB extraction was performed for samples obtained from the above cultures and was analyzed using the GC-MS. The fraction of carbon derived from glucose for isobutanol (F_{IB}) was estimated based on the isobutanol MS peak abundances:

$$F_{IB} = \frac{\sum_{i=1}^4 i \times A_i}{4 \sum_{i=0}^4 A_i} \quad (4.2)$$

where A_i is the abundance of the mass-to-charge ratio peaks for the various isobutanol isotopomers (i.e., $A_0 \sim A_4$ for $m/z=74 \sim 78$).

4.3 Results

4.3.1 Construction of an isobutanol producing *Synechocystis* 6803 strain.

IB synthesis in *Synechocystis* 6803 requires the expression of two heterologous genes of the Ehrlich pathway. The enzymes 2-keto-acid decarboxylase and alcohol dehydrogenase can convert 2-keto acids into alcohols. In this work, we constructed a plasmid pTKA3 containing the genes *kivd* and *adhA* from *Lactococcus lactis* under the control of an IPTG inducible promoter, P_{tac} . The plasmid was designed to integrate the genes into a neutral site in the genome of *Synechocystis* 6803, along with a kanamycin resistance cassette (Fig. 4.1a - Left). The wild-type strain of *Synechocystis* 6803 was transformed with pTKA3, resulting in the recombinant strain AV03. The integration of the insert genes into the genome was verified by a colony PCR after several rounds of segregation (Fig. 4.1a - Right).

To identify the optimal IPTG concentration required for IB synthesis, the AV03 strain was grown under different concentrations of IPTG. IB analysis from the cultures indicated that IB was highly synthesized even without the addition of IPTG (Fig. 4.1b). To verify if this observation was an artifact of any mutations that might have occurred in *lacI* or the promoter, the foreign DNA integrated into the chromosome of AV03 was sequenced. Sequencing results for the *lacI* and the promoter P_{tac} in the genome of AV03 revealed that the nucleotide sequence was completely intact. There have been reports of leaky expressions with IPTG inducible promoters¹⁶³. Besides, Fig. 4.1b indicates that as the concentration of IPTG went higher than 1mM, the IB synthesis reduced. The OD_{730} of the different cultures indicated that the addition of IPTG did

not apparently interfere with the growth rate of the culture. RT-PCR showed the expression levels of the genes *kivd* and *adhA* under different IPTG concentrations. The result of RT-PCR experiment (Fig. 4.1b - Inset) indicated that the levels of *kivd* and *adhA* mRNA synthesized in the mutant were higher with IPTG than without. Henceforth, lower expression of the two genes is sufficient for IB synthesis, possibly because the Ehrlich pathway may not be the rate-limiting step for IB production.

4.3.2 Isobutanol synthesis under autotrophic and mixotrophic growth.

Under autotrophic conditions, *Synechocystis* 6803 utilizes light as an energy source (ATP and NADPH) for the conversion of CO₂ into biomass and IB. Fig. 4.2a compares the autotrophic growth of the mutant and the wild strain. Under autotrophic conditions, we found that the growth rate of the mutant AV03 remains unaltered as compared to the wild-type strain. IB accumulation in the mutant was tested under autotrophic condition (Fig. 4.2b), and the strain was found to synthesize a maximum of 90 mg/L of IB (the only extracellular product detected by GC-MS) in a 6-day culture. In a sealed shaking flask, NaHCO₃ in the medium (50mM) was consumed by AV03 within six days, and then both the biomass and IB started declining.

The wild-type strain of *Synechocystis* 6803 grows about 5 times faster under mixotrophic conditions compared to autotrophic conditions (Fig. 4.2a). However, our mutant AV03 did not exhibit an increased growth rate under mixotrophic conditions. To measure the glucose utilization by wild-type and mutant AV03, we fed cells with 0.5% fully labeled glucose and nonlabeled bicarbonate. Isotopomer analysis of ¹³C-abundance in cell metabolites (Fig. 4.2c) showed that the wild-type synthesized 70~90% of its amino acids using carbons from glucose, whereas the mutant produced biomass only using 5~10% carbon from glucose, and 12% of the carbon of IB was labeled (i.e., derived from glucose). These results indicated that the mutant

tended to limit glucose metabolism for IB production. The AV03 strain was found to synthesize a maximum of 114 mg/L of IB mixotrophically after 9 days (Fig. 4.2b), whereas cells with only glucose (heterotrophic without bicarbonate or CO₂) synthesized a maximum of 27 mg/L of IB. This result suggests that the *Synechocystis* 6803 mutant is unable to take significant advantage of its glucose metabolism to have a fast rate of IB production.

4.3.3 In situ alcohol concentrating system using a solvent trap.

IB is toxic to the cells and our study revealed that IB inhibited *Synechocystis* 6803 growth at external concentrations of only 2 g/L (Fig. 4.3). Moreover, our control experiments indicated the loss of IB after 9 days of continuous incubation. IB can be slowly degraded by photochemically-produced hydroxyl radicals in aerobic cyanobacterial cultures¹⁶⁴⁻¹⁶⁶. Therefore, a system with continuous removal of the synthesized alcohol products will be beneficial^{167, 168}. We have demonstrated the use of an *in situ* alcohol removal system by using oleyl alcohol^{169, 170} as a solvent trap for increasing IB titer. In previous studies, gas stripping is one efficient way for IB recovery, but it requires an expensive cooling system due to very low concentrations of IB from photo-bioreactors. Here, inside each cultivation flask, we placed a small glass vial containing 0.5 or 1 mL oleyl alcohol solvent, so that oleyl alcohol was not mixed with the culture solution (Fig. 4.4). Volatile IB in the headspace can be trapped in the solvent vial because of the high solubility of IB in oleyl alcohol. This method will effectively trap the IB while the solvent will not directly interfere with light and cell culture conditions.

To test the effect of oleyl alcohol on IB productivity, we did a 3-week time-course study (Fig. 4.4) by adding 50 mM NaHCO₃ intermittently (every 4 days). During the cultivation, the pH of the cultures was also adjusted to be between 8 and 9 before the addition of excess bicarbonate. Every 3 days, the oleyl alcohol in the vials was taken out for IB measurement, and

then replaced with fresh oleyl alcohol. The mixotrophic cultures with alcohol trap (0.5 mL) reached the highest net IB concentration of 298 mg/L. The autotrophic cultures with 0.5 mL and 1 mL oleyl alcohol had a maximum net IB titer of 180 mg/L and 240 mg/L, respectively, whereas the autotrophic cultures without the oleyl alcohol trap were able to achieve a maximum of only 108 mg/L of IB. IB levels in the organic phase reached concentrations of up to 500 mg/L with only 3 days of trapping.

4.4 Discussion

Isobutanol (IB) is a promising biofuel for the replacement of gasoline. So far, *E. coli* has remained the most successful microbial host for IB production. In this study, we have focused our efforts on a cyanobacterial species, *Synechocystis* 6803, which can grow on both CO₂ and glucose. The mixotrophic cultivation may offer industrial flexibility and economic benefits because the gas-liquid mass transfer of CO₂ is often a rate-limiting step in efficient photobioreactor operations. Attempts in creating a stable strain of *Synechococcus* 7942 that can transport and utilize glucose has been barely successful¹⁷¹. The glucose tolerant strain of *Synechocystis* 6803 unlike other cyanobacterial strains, can perform both autotrophic and mixotrophic metabolisms. In our work, we found that the wild-type strain of *Synechocystis* 6803 under mixotrophic conditions grew at a rate 5 times faster than the autotrophic condition. Moreover, the engineered *Synechocystis* 6803 strain accumulated 90 mg/L of IB, whereas the *Synechococcus* 7942 strain expressing the same two enzymes (keto-acid decarboxylase and alcohol dehydrogenase) only accumulated a maximum of 18 mg/L²⁸. Switching the condition from autotrophic to mixotrophic for the mutant AV03 increased the maximum IB titer to 114

mg/L. Interestingly, the mutant tended to grow autotrophically and had minimal glucose utilization compared to the wild-type strain (Fig 4.2c).

IB can be inhibitory to cell physiologies. Moreover, our experiments also observed IB degradation (by hydroxyl radicals) during the incubation process. Thereby, efforts in coming up with product recovery are important to improve IB titer in cyanobacterial culture. This work employed an *in situ* IB removal system by growing cultures in shake flasks with vials containing oleyl alcohol. Mixotrophic growth of AV03 along with *in situ* IB removal synthesized a maximum of 298 mg/L IB, which is lower than the highest IB titer (450 mg/L) reported in *Synechococcus* 7942 mutant expressing 3 more genes of the keto acid pathway. On the other hand, our strain design has two apparent advantages for industrial application. First, our strain does not require any antibiotics to maintain its IB production because the two heterologous genes in the mutant show good stability during normal cultivation conditions. Second, the strain does not need any inducer (IPTG) for IB production, which can significantly reduce the industrial costs.

Overexpressing the keto acid pathway can increase the IB titer in *Synechococcus* 7942²⁸. Furthermore, optimizing CO₂ and light conditions of the cyanobacterial strain can also increase the final titer and productivity. Liu et. al.,¹⁷² have reported a doubling time of 7.4 hours for *Synechocystis* 6803, by growing them under 140 μmol of photons m⁻² s⁻¹ of light and by bubbling 1% CO₂ enriched air. Therefore, our strain can serve as a springboard for future development of higher performance *Synechocystis* 6803 strains with increased IB titer and productivity.

In summary, IB synthesis under autotrophic conditions in a cyanobacterium *Synechocystis* 6803 was demonstrated by the expression of two heterologous genes. It was further demonstrated that mixotrophic cultures of the mutant can significantly increase IB synthesis with minimal glucose consumption. The mechanism behind the reduced glucose-utilizing metabolism of AV03 compared to the wild-type strain remains unclear. A possible explanation is that the cells tend to avoid the intracellular metabolic imbalance or IB intermediate inhibition by down-regulating glucose uptake. Using oleyl alcohol as a simple solvent trap, IB production can be improved by 2~3 times. Therefore, *in situ* IB recovery may reduce the product loss and separation cost. We have also demonstrated that a simple expression of the Ehrlich pathway with bioprocess modification can synthesize IB without other major waste products, while still achieving comparable levels of IB to an extensive genetically modified *Synechococcus* 7942 strain (Table 4.1) ²⁸.

Acknowledgment

The authors wish to thank Dr. Abhay Singh and Dr. Anindita Bandyopadhyay for their technical guidance. The authors would also like to thank Amelia Nguyen and Amelia Chen for their assistance with data collection for a few experiments. This research was funded by an NSF Career Grant to YT (MCB0954016), a grant to HBP from the Office of Science (BER), U.S. Department of Energy, and a grant to HBP from the Consortium for Clean Coal Utilization at Washington University. We would also like to thank Seema Mukhi Dahlheimer from the Engineering Communication Center of Washington University in St. Louis for her close reading of this manuscript.

Figure 4.1a

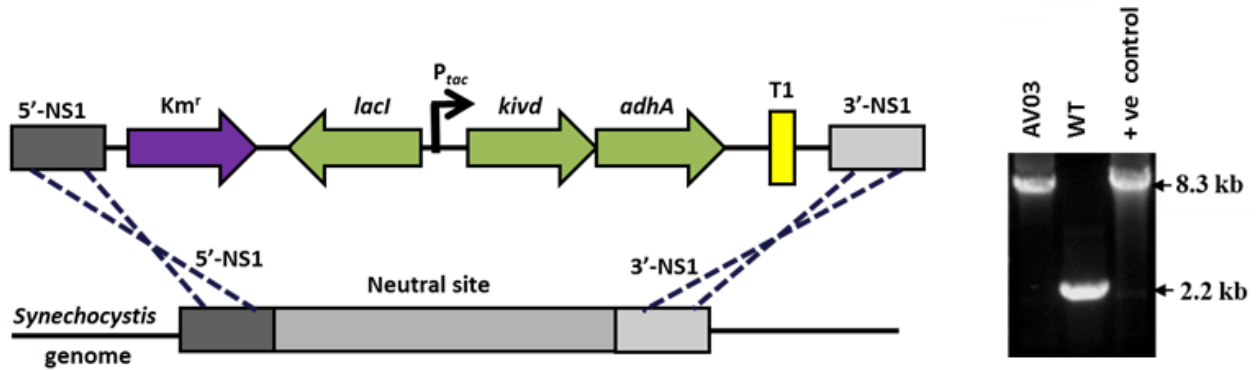


Figure 4.1b

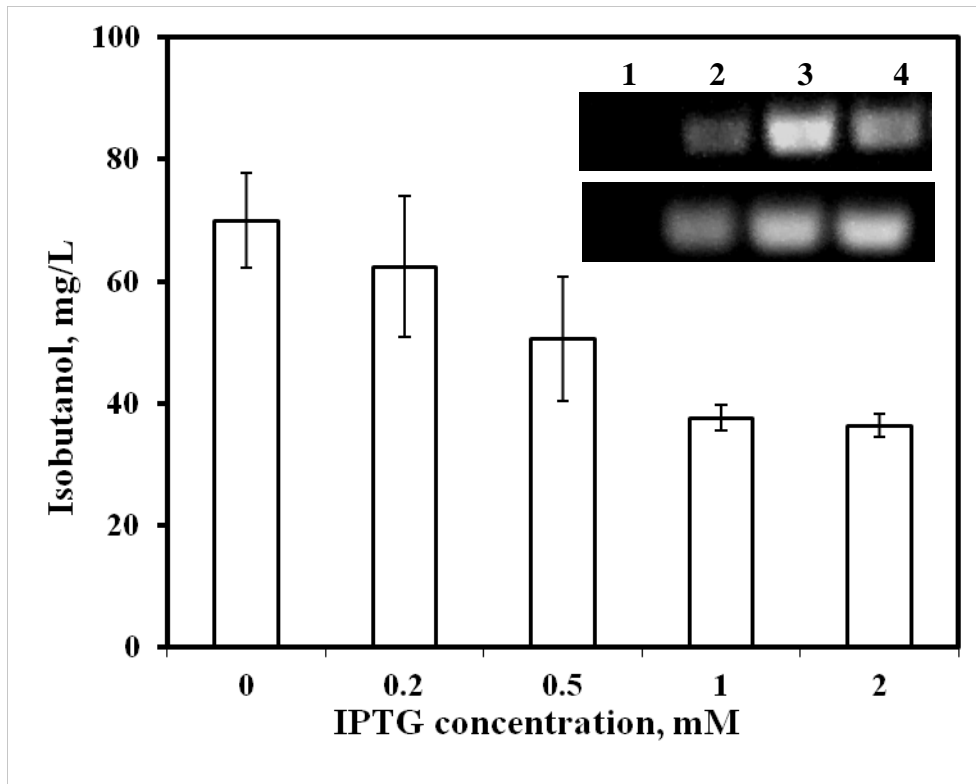


Figure 4.1: (a) Schematic representation to show the integration of the genes *kivd* and *adhA* into the genome of *Synechocystis* 6803. Colony PCR performed to verify the integration of the insert into the genomic DNA of the mutant (AV03). The vector *ptka3* w was used as a template

for the positive control and wild-type cells were used as negative control. Colony PCR of AV03, showed the presence of a band (8.3kb) the same size as the positive control (+ve) and the absence of the negative control (WT) band. (b) IB synthesized by engineered *Synechocystis* 6803 under different IPTG concentrations (n=3). (Inset) Result of an RT-PCR performed to detect the expression of the heterologous genes *kivd* (Top: 500bp from *kivd*) and *adhA* (Bottom: 200bp from *adhA*). Lane 1, wild-type 6803 (WT); Lane 2, AV03 with 0 mM IPTG; Lane 3, AV03 with 0.5mM IPTG; Lane 4, AV03 with 1mM IPTG.

Figure4. 2a

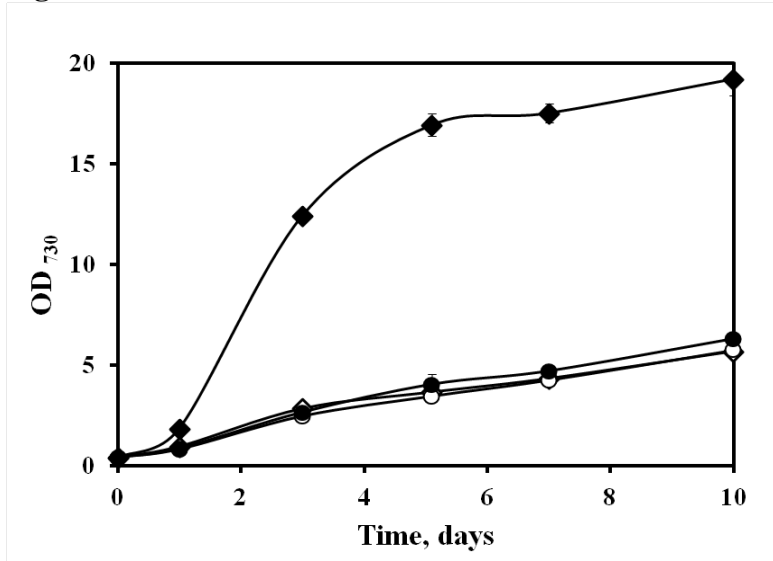


Figure4. 2b

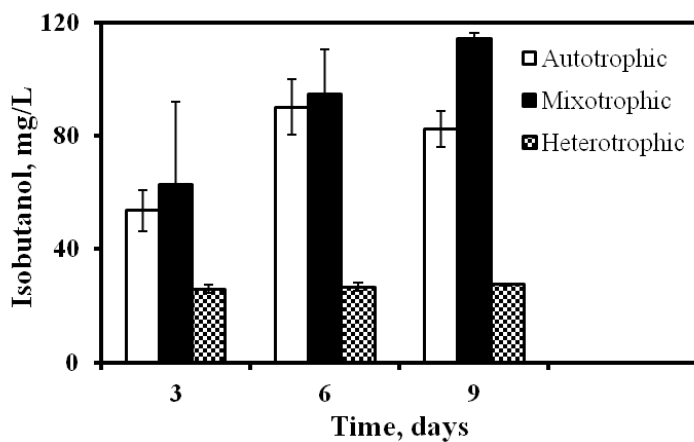


Figure 4.2c

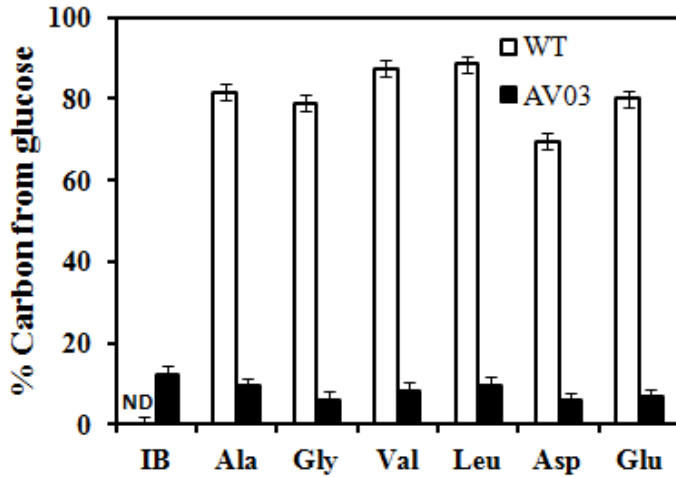


Figure 4.2: (a) Growth curves of *Synechocystis* 6803 WT and AV03 (n=3, shake flask cultures): \diamond WT under autotrophic, \blacklozenge WT under mixotrophic, \circ AV03 under autotrophic and \bullet AV03 under mixotrophic conditions (note: growth curve of AV03 under mixotrophic condition overlaps with autotrophic growth curves of AV03 and WT). (b) IB synthesized in AV03 under autotrophic conditions (only HCO_3^-), heterotrophic (only glucose) and mixotrophic (both HCO_3^- and glucose) conditions (n=3, shake flask cultures with closed caps). (c) Percentage carbon contribution of glucose for synthesizing amino acids and isobutanol in the wild-type (WT) and the mutant strain (AV03) as measured on day 9 (shake flask cultures with closed caps). Isotopomer analysis (TBDMS based method) of proteinogenic amino acids confirms the low ^{13}C -glucose utilization by the mutant. The error bar represents the 2% technical error of the instrument.

Figure 4.3

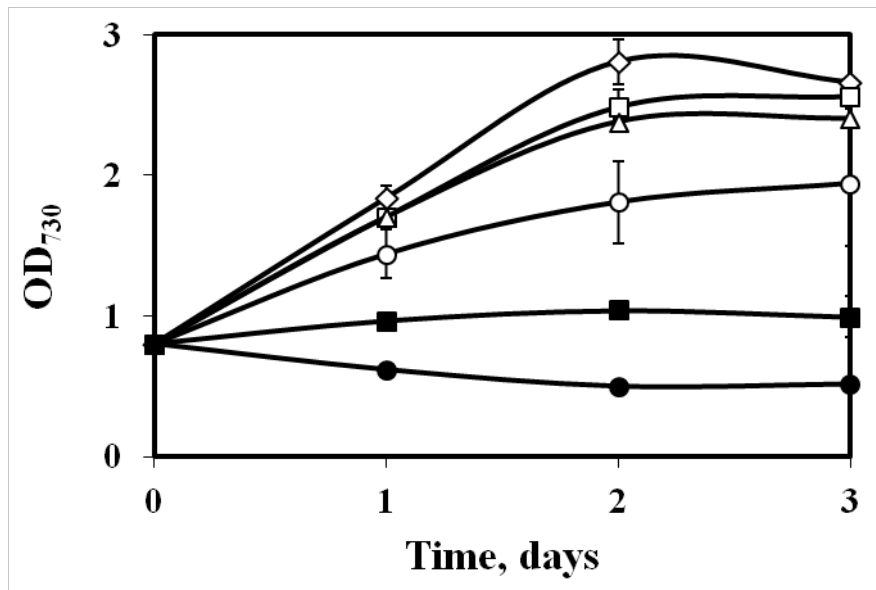


Figure 4.3: Toxic effects of IB on the growth of *Synechocystis* 6803. IB was added to a final concentration (g/L, n=2) of \diamond 0, \square 0.2, \triangle 0.5, \circ 1, \blacksquare 2 and \bullet 5 to a *Synechocystis* 6803 culture with an initial OD₇₃₀ ~ 0.8.

Figure 4.4

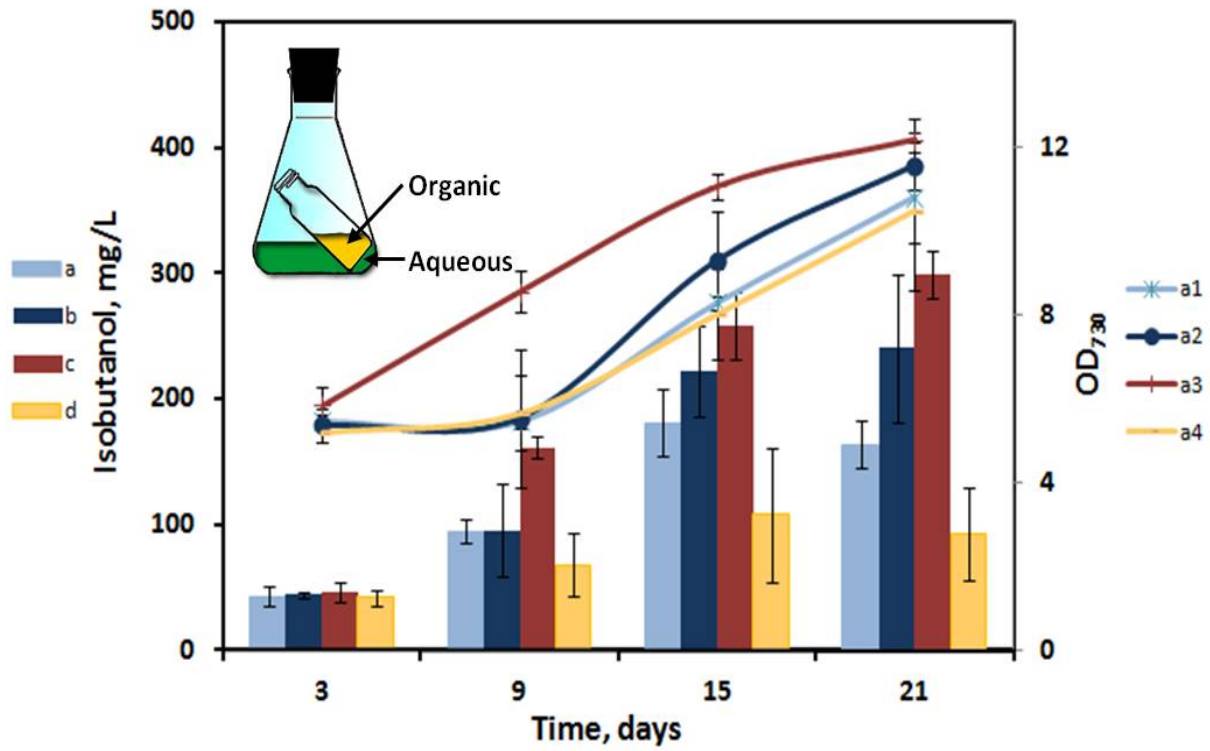


Figure 4.4: Net concentration of IB synthesized (columns) and biomass growth (curves) by the AV03 culture under different conditions (n=3). a – IB with 0.5 mL oleyl alcohol (Autotrophic); b – IB with 1 mL oleyl alcohol (Autotrophic); c – IB with 0.5 mL oleyl alcohol and glucose (Mixotrophic); d – IB with no oleyl alcohol (Autotrophic, negative control); a1 – OD₇₃₀ with 0.5 mL oleyl alcohol (Autotrophic); a2 – OD₇₃₀ with 1 mL oleyl alcohol (Autotrophic); a3 – OD₇₃₀ with 0.5 mL oleyl alcohol and glucose (Mixotrophic); a4 – OD₇₃₀ with no oleyl alcohol (Autotrophic, negative control). (**Inset**) Schematic representation of the *in situ* IB removal system used to increase the production of IB.

Table 4.1: Metabolic engineering of cyanobacterial strains for biofuel production.

Product	Species	Titer or Productivity	Overexpressed genes	Promoters	Culture vessel / Remarks	Culture Days	Ref.
Ethanol	<i>Synechococcus</i> 7942	230 mg/L	<i>pdc</i> and <i>adh</i>	<i>rbcLS</i>	Shake flask	28 days	¹⁷³
Ethanol	<i>Synechocystis</i> 6803	552 mg/L	<i>pdc</i> and <i>adh</i>	<i>psbA2</i>	Photobioreactor	6 days	²¹
Isobutyraldehyde	<i>Synechococcus</i> 7942	1100 mg/L	<i>alsS</i> , <i>ilvC</i> , <i>ilvD</i> , <i>kivd</i> and <i>rbcls</i>	<i>LlacO₁</i> , <i>trc</i> and <i>tac</i>	Roux culture bottle with NaHCO ₃	8 days	²⁸
Isobutanol	<i>Synechococcus</i> 7942	18 mg/L	<i>kivd</i> and <i>yqhD</i>	<i>trc</i>	Shake flask with NaHCO ₃	1 day	²⁸
Isobutanol	<i>Synechococcus</i> 7942	450 mg/L	<i>alsS</i> , <i>ilvC</i> , <i>ilvD</i> , <i>kivd</i> and <i>yqhD</i>	<i>LlacO₁</i> , <i>trc</i>	Shake flask with NaHCO ₃	6 days	²⁸
Fatty alcohol	<i>Synechocystis</i> 6803	200±8 µg/L	<i>far</i>	<i>rbc</i>	Photobioreactor with 5% CO ₂	18 days	³⁰
Alkanes	<i>Synechocystis</i> 6803	162±10 µg/OD/L	<i>accBCDA</i>	<i>rbcl</i>	Shake flask	-	³⁰
Fatty acids	<i>Synechocystis</i> 6803	197 ±14 mg/L	<i>tesA</i> , <i>accBCDA</i> , <i>fatB1</i> , <i>fatB2</i> , <i>tesA137</i>	<i>psbA2</i> , <i>cpc</i> , and <i>trc</i>	1% CO ₂ bubbling	17 days	¹⁷²
Hydrogen	<i>Synechococcus</i> 7942	2.8 µmol/hr/mg Chl-a*	<i>hydEF</i> , <i>hydG</i> and <i>hydA</i>	<i>psbA1</i> , <i>lac</i>	Anaerobic conditions with DCMU treatment**	-	³²
1-Butanol	<i>Synechococcus</i> 7942	14.5 mg/L	<i>hbd</i> , <i>crt</i> , <i>adhE2</i> , <i>ter</i> and <i>atoB</i>	<i>trc</i> , <i>LlacO₁</i>	Dark roux culture bottle under anoxic condition	7 days	²⁹
Fatty alcohol	<i>Synechocystis</i> 6803	20 ± 2 µg/L/OD	<i>far</i> , <i>aas</i>	<i>rbc</i> , <i>psbA2</i>	Shake flask	-	¹⁷⁴
1-Butanol	<i>Synechococcus</i> 7942	30 mg/L	<i>ter</i> , <i>nphT7</i> , <i>bldh</i> , <i>yqhD</i> , <i>phaJ</i> , <i>phaB</i>	<i>trc</i> , <i>LlacO₁</i>	Shake flask	18 days	¹⁷⁵

where * Chl-a is chlorophyll a; **DCMU is 3-(3,4-dichlorophenyl)-1,1-dimethylurea.

Table 4.2: Primer sequences used in this study.

Name	Sequence (5'→3')
AMV14F	GCGCACATGTCGGAACAGGACCAAGCCTTGAT
AMV15R	GCGC CCTGAGGCCTTTACCATGACCTGCAGGG
AMV16F	GCGCGACGGGGAGTCAATTGTGCCATTGCCATAACTGCTTTCG
AMV17R	GCGCGACTCCCCGTCTTTGACTATCCTTTTTAGGATGGGGCA
ps1_up_fwd	TACCGGAACAGGACCAAGCCTT
AMV01	GCGCCATATGTATACAGTAGGAGATTACCTATTAGAC
AMV12	GCAGCAGCAACATCAACTGGTAAG
AdhA-TMs	TCAACTAGTGGTACCAGGAGATATAATATGAAAGCAGCAGTAGTAAG
adhA_RTr	GACAATTCCAATTCCTTCATGACCAAG
Rnpbr	CGGTATTTTTCTGTGGCACTGTCC
Rnpbf	CAGCGGCCTATGGCTCTAATC
AV03_6F	GAATCCGTAATCATGGTCATAGCTG
AV03_4R	GCCAAAGCTAATTATTCATGTCCTGT
AV03_1R	TGTCGGGGCGCAGCCATGA
AV03_2F	AGAGGATCCTTCTGAAATGAGCTG
AV03_3F	CAGAGCCTAATCTTAAAGAATTCGTGG
AV03_4F	GGGTAAACTATTTGCTGAACAAAATAAATC
AV03_2R	CCGCTTCTGCGTTCTGATTTAATC
AV03_5F	GTTGATCGGCGCGAGATTTAATCG
AV03_7F	CCGTTGAAATTGACCGAGTACTTTCT
AV03_8F	CAGTCGAAAGAGAAATTCATGGACC
AV03_5R	CGCTACGGCGTTTCACTTCTG

Chapter 5: Photoautotrophic production of D-lactic acid in an engineered cyanobacterium

This chapter has been reproduced from the following publication:

Varman, A.M., Yu, Y., You, L., & Tang, Y.J. Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microbial Cell Factories*, 12, 117 (2013).

AMV and YY contributed equally for this work.

Abstract

Background: The world faces the challenge to develop sustainable technologies to replace thousands of products that have been generated from fossil fuels. Due to concerns about food security, sugar-based microbial fermentation raises economical questions. Thus, phototrophic microbial cell factories serve as promising alternatives for the production of commodity chemicals and biofuels. For example, polylactic acid (PLA) with its biodegradable properties is a sustainable, environmentally friendly alternative to polyethylene. At present, PLA microbial production is mainly dependent on food crops such as corn and sugarcane. Moreover, optically pure isomers of lactic acid are required for the production of PLA, where D-lactic acid controls the thermochemical and physical properties of PLA. Henceforth, production of D-lactic acid through a more sustainable source (CO₂) is desirable.

Results: We have performed metabolic engineering on the cyanobacterium, *Synechocystis* sp. PCC 6803, for the phototrophic synthesis of optically pure D-lactic acid from CO₂ by utilizing solar energy. Synthesis of optically pure D-lactic acid was achieved by utilizing a recently discovered enzyme, (i.e., a mutated glycerol dehydrogenase, GlyDH*). Significant improvements in D-lactate synthesis were achieved through codon optimization and by

balancing the cofactor (NADH) availability through the heterologous expression of a soluble transhydrogenase (STH). We have also discovered that addition of acetate to the cultures improved lactic acid production. More interestingly, ¹³C based metabolic pathway analysis revealed that acetate was not used for the synthesis of lactic acid, but was mainly used for synthesis of some biomass building blocks (such as leucine and glutamate). Finally, the optimal strain was able to accumulate 1.14 g/L (photoautotrophic condition) and 2.17 g/L (phototrophic condition with acetate) of D-lactate in 24 days.

Conclusions: We have demonstrated the photoautotrophic production of D-lactic acid by engineering a cyanobacterium, *Synechocystis* 6803. The engineered strain shows an excellent D-lactate productivity from CO₂. In the late growth phase, the lactate production rate by the engineered strain reached a maximum of 0.19 g D-lactate/L/day (in the presence of acetate). This study serves as a good complement to the recent engineering work done on *Synechocystis* 6803 for L-lactate production. Thereby, our study may facilitate developments in the use of cyanobacterial cell factories for the commercial production of high quality PLA.

5.1 Background

Fossil fuels helped literally ignite the industrial revolution, and from then on radically changed the way we live; today, thousands of products are generated from fossil fuels¹⁷⁶. Unfortunately, fossil fuels are non-renewable and their reserves will foreseeably run dry. Moreover, the reckless use of this resource has resulted in a tremendous release of greenhouse gases leading to adverse effects to our earth's climate and to the creatures living on our planet. These drawbacks have driven researchers to look for alternative renewable replacements for petroleum and petroleum-derived products. Amongst the petroleum-derived products; polyethylene with an annual productivity of 80 million metric tons per annum stands out as one of the most commonly used plastics¹⁷⁷. Polylactic acid (PLA) is made by the polymerization of lactic acid and has the potential to replace polyethylene as a biodegradable alternative¹⁷⁸. Lactic acid is a chiral compound and exists in two isomeric forms: D (-) lactic acid and L (+) lactic acid. The various properties of polylactic acid are modulated by the mixing ratio of the D (-) and L (+) lactic acid and, henceforth, it is essential to produce both the isomers¹⁷⁹. It has been estimated that for the PLA production to be profitable, the lactic acid price should be less than 0.8\$/kg¹⁸⁰. This necessitates the production of lactic acid from a cheaper source. Although microbial fermentation can produce lactate from sugar-based feedstock, such process may compete with global food supplies. Therefore, this work focuses on cyanobacterial process development for the sustainable synthesis of D (-) lactic acid, with CO₂ as the carbon substrate and sunlight as an energy source.

Cyanobacteria have the ability to reduce atmospheric CO₂ into useful organic compounds by using solar energy and have been engineered to synthesize a number of value-

added products^{28, 29, 181, 182}. *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803) with its ability to uptake foreign DNA naturally, has been the model organism of choice for various metabolic engineering works^{135, 183, 184}. *Synechocystis* 6803 also has the ability to grow mixotrophically with glucose and acetate¹⁸⁵. Therefore, along with CO₂, its versatile carbon metabolism allows the co-utilization of cheap organic compounds for product biosynthesis. For example, acetate abundant wastewater generated from biomass hydrolysis and anaerobic digestion¹³¹, can be potentially used for promoting cyanobacterial productivity. More importantly, there are numerous molecular biology tools for *Synechocystis* 6803, making it an attractive organism for metabolic engineering works^{163, 186}.

Synechocystis 6803 has recently been engineered for the production of L-lactate (a maximal titer of 1.8 g/L and a maximal productivity of 0.15 g/L/day)¹⁸⁷⁻¹⁸⁹. However, engineering *Synechocystis* 6803 for the production of optically pure D-lactate synthesis is more difficult due to the lack of an efficient D-lactate dehydrogenase. Recently, a mutated glycerol dehydrogenase (GlyDH*) was discovered by Wang et al.¹⁹⁰ and this enzyme was found to behave as a D-lactate dehydrogenase, exhibiting an unusually high specific activity of 6.9 units per mg protein with pyruvate and NADH as substrates. This enzyme allows a *Bacillus coagulans* strain to produce 90g/L of D-lactate. Their work served as a motivation for us to engineer *Synechocystis* 6803 through the heterologous expression of *gldA101* (encodes GlyDH*). We found that this original enzyme was able to synthesize optically pure D-lactate in *Synechocystis* 6803. To further improve cyanobacterial productivity, we employed three strategies: 1. Codon optimization of *gldA101* (Supplementary Figure 5.5); 2. Heterologous expression of a transhydrogenase; 3. Supplementing cultures with extracellular carbon sources (such as glucose,

pyruvate and acetate). The final engineered strain demonstrated a high D-lactic acid productivity and titer (titer >1g/L).

5.2 Results and Discussion

Cyanobacteria need a lactate dehydrogenase to synthesize lactate from pyruvate (Figure 5.1). Earlier works on *Synechocystis* 6803 for lactate production involved the expression of an *ldh* from *Bacillus subtilis* for synthesis of L-lactate¹⁸⁸. As a first step, we tested the activity of GlyDH* for D-lactate production¹⁹⁰ by transferring the gene from *Bacillus coagulans* to *Synechocystis* 6803. A plasmid pYY1 was constructed that contained the gene *gldA101* under the control of an Isopropyl β -D-1-thiogalactopyranoside (IPTG) inducible promoter, P_{trc}. The *gldA101* gene was then subsequently transferred to the glucose tolerant wild type *Synechocystis* 6803 through natural transformation, generating the strain AV08. The optical density and the D-lactate concentration of the AV08 cultures were monitored in shake flasks. As can be verified from Figure 5.2, AV08 did not show any significant levels of D-lactate in the initial 12 days. The D-lactate levels started increasing steadily at the late autotrophic growth phase and reached a final titer of 0.4 g/L, whereas a wild type strain of *Synechococcus* 7002 was able to produce only ~ 7 mg/L of D-lactate through glucose fermentation¹⁹¹.

A familiar strategy to increase the synthesis of a target product would be to increase the levels of the heterologous enzyme inside the cell. This can be achieved by modifying the enzyme regulation either at the transcriptional level or at the translational level. Cyanobacteria are known to have their own preference in the use of codons for synthesizing amino acids¹⁹². Lindberg et al.¹⁹³ have employed codon optimization for the isoprene synthase gene *IspS* and have found a 10-fold increase in the *IspS* expression level. More recently, this strategy was

applied to increase the expression of the *efe* gene (from *Pseudomonas syringae*) in *Synechocystis* 6803 for ethylene production¹⁹⁴. Since the gene involved in this work was borrowed from a gram-positive organism and *Synechocystis* 6803 being gram-negative, we hypothesized that this would be a useful strategy. The codon optimized gene *gldA101-syn* (synthesized by Genewiz Inc, South Plainfield, NJ) was integrated into the *psbA1* gene loci in the genome of the WT *Synechocystis* 6803 using the plasmid pDY3 to obtain the strain AV11.

Further improvements in product synthesis can be achieved by rectification of bottlenecks in the metabolic pathway. The lactate dehydrogenase enzyme utilizes NADH as its cofactor, whereas the ratio of NADH to NADPH is reported to be much lower in cyanobacteria. For example, the ratio of NADH to NADPH in *Synechococcus* 7942 under light conditions was estimated to be 0.15, and in *Synechocystis* 6803 under photoautotrophic conditions the intracellular NADH concentration was only 20 nmol/g fresh weight, whereas the intracellular concentration of NADPH was about 140 nmol/g fresh weight¹⁹⁵⁻¹⁹⁷. This lower concentration of NADH in cyanobacteria, points to the fact that availability of NADH could be a major limiting factor for synthesizing D-lactate. Henceforth, a soluble transhydrogenase, *sth* from *Pseudomonas aeruginosa*¹⁹⁸, was introduced downstream of the gene *gldA101-syn*. This engineered strain was called AV10. The heterologous genes in AV10 and AV11 are under the control of the same single promoter, P_{trc}, located upstream of *gldA101-syn* and *sth* in AV10 and located upstream of *gldA101-syn* in AV11.

The three strains (AV08, AV10 and AV11) showed similar growth rates to wild type strain under photoautotrophic conditions, and thus the production of D-lactate did not introduce growth defects in the engineered strains (Figure 5.2A and Figure 5.6). However, the three strains

differed in the production rate of D-lactic acid. The strain AV11 with codon optimization (*gldA101-syn*) had an improved productivity for D-lactate compared to the AV08 strain (Figure 5.2B). Both strains produced D-lactate mainly during the later growth stage. Introduction of the transhydrogenase improved the D-lactate synthesis further in AV10, and this strain produced D-lactate in both the growth phase and non-growth phase. The rate of photoautotrophic D-lactate production by AV10 increased significantly (achieving a maximum productivity of 0.1 g/L/day and ~0.2 mmol/g cell/day) during the late phase of the culture and the final titer of D-lactate reached 1.14 g/L.

We observed that the D-lactate production rate reached its peak in the later stages of cultivation, suggesting that more carbon flux has been directed to lactate production during the non-growth phase. This increased flux was expected because the lactate precursor (pyruvate) is a key metabolic node occupying a central position in the synthesis of diverse biomass components, and more pyruvate becomes available for lactate synthesis when biomass growth becomes slow. Therefore, an obvious thought would be to enhance lactate production by supplementing the cultures with pyruvate¹⁹⁹. However, our experiments found that addition of pyruvate did not yield apparent improvements in D-lactate synthesis (data not shown), possibly because *Synechocystis* 6803 may lack an effective pyruvate transporter. The alternate option would be to grow AV10 with glucose and increase the glycolysis flux for pyruvate synthesis. In our previous study, addition of glucose was found to increase isobutanol production in *Synechocystis* 6803¹³⁶. However in this study, when we grew the AV10 strain under mixotrophic conditions (with 5 g/L glucose), it did not show a higher growth rate or display improvements in the final D-lactate titer compared to the autotrophic condition. The AV10 cultures grown in the presence of glucose

instead showed an impaired growth, possibly because the engineered pathways caused a metabolic imbalance during glucose catabolism (Figure 5.3).

We also hypothesized that the intracellular pyruvate pool can be increased for lactate production by addition of exogenous acetate. Supplementing cultures with acetate can redirect more carbon from pyruvate to lactate in three possible ways²⁰⁰: (1) acetate is used as a building block for lactate production; (2) acetate provides additional carbon source for biomass synthesis and reduce pyruvate consumption; (3) acetate conversion by acetyl-CoA synthetase consumes Coenzyme-A (CoA), decreasing the CoA pool available for pyruvate decarboxylation. To test this hypothesis, the AV10 cultures were supplemented with 15mM acetate. We found that growth rate of the AV10 cultures with acetate (Figure 5.3A) remained comparable to their growth rate under autotrophic condition, but there was substantial improvement in the synthesis of D-lactate (the maximal titer reached 2.17 g/L and the peak productivity reached ~0.19 g/L/day, Figure 5.3B).

To further understand the role played by glucose and acetate in D-lactate synthesis, AV10 cultures were grown with [1,2-¹³C] glucose and [1,2-¹³C] acetate (Sigma, St. Louis). Cultures were collected from the mid-log phase and were used for amino acid and D-lactate analysis. As an example, mass spectrum of D-lactate from a cyanobacterial culture is shown in supplementary Figure 5.7. The ¹³C abundance in the amino acids and lactate were obtained as mass fraction m_i , where 'i' indicates the number of ¹³C in the molecule. As can be seen from Figure 5.4A, glucose-fed cells have significant ¹³C-carbon distributed in amino acids (indicated by an increase in m_1 and m_2). Also, D-lactate from glucose-fed cultures was partially ¹³C-labeled (m_2 ~0.22). The isotopomer data in Figure 5.4A proved that ¹³C-glucose provided the carbon source for both biomass and lactate production. However, glucose-based mixotrophic fermentation is not

beneficial to D-lactate production compared to autotrophic cultures, possibly because carbon flux from glycolysis may cause some carbon and energy imbalance¹³⁶. As for the acetate-fed cultures, only leucine and glutamate (which both use acetyl-CoA as their precursor) were significantly labeled (an m_2 of 0.31 and 0.32 respectively), while other amino acids (e.g., aspartate and alanine) were nonlabeled (Figure 5.4B). Interestingly, D-lactate from acetate-fed culture was almost nonlabeled, indicating that the carbons of lactate molecules were mainly derived from CO₂. Therefore, the observed enhancement of lactate synthesis in the presence of acetate can be explained by two complementary mechanisms. First, acetate is an additional carbon source for synthesizing biomass building blocks, such as fatty acids and some amino acids, thus redirecting the extra carbon flux from CO₂ to lactate. Secondly, acetate may limit the pyruvate decarboxylation reaction by reducing the CoA pool by the formation of acetyl-CoA and thus improve pyruvate availability for lactate synthesis.

5.3 Conclusions

The results reported here are for the autotrophic production of D-lactate in cyanobacteria via the heterologous expression of a novel D-lactate dehydrogenase (GlyDH*) and by balancing the precursors and cofactors. Other molecular strategies may also be applied to further improve the D-lactate production: (1) by seeking stronger promoters¹⁸⁶; (2) optimizing ribosomal binding sites²⁰¹; (3) improving activity of GlyDH* via protein engineering; (4) introducing powerful lactate transporter²⁰²; (5) knocking out competing pathways (such as the glycogen and polyhydroxybutyrate synthesizing pathways); (6) duplicating the heterologous genes by integrating at multiple sites²⁰³; and (7) limiting biomass production by knocking down the pyruvate decarboxylation reaction. Also, considering the future outdoor algal processes for

scaled up D-lactate production, we hypothesize that knocking out metabolic pathways that synthesize carbon storage molecules (polyhydroxybutyrate and glycogen) may be deleterious to algal growth during the night phase in day-night cultivation²⁰⁴. On the other hand, process optimization by employing better light conditions, along with proper CO₂ concentration, pH and temperature control, may also be employed to increase the D-lactate productivity in a scaled-up system.

5.4 Materials and methods

5.4.1 Chemicals and reagents. Restriction enzymes, Phusion DNA polymerase, T4 DNA ligase and 10-Beta electro-competent *E. coli* kit were purchased from Fermentas or New England BioLabs. Oligonucleotides were purchased from Integrated DNA Technologies (IDT). All organic solvents, chemicals, ¹³C-labeled acetate, and glucose used in this study were purchased from Sigma-Aldrich (St. Louis, MO).

5.4.2 Medium and growth conditions. *E. coli* strain 10-Beta was used as the host for all plasmids constructed in this study. *E. coli* cells were grown in liquid Luria-Bertani (LB) medium at 37°C in a shaker at 200 rpm or on solidified LB plates. Ampicillin (100 µg/mL) or kanamycin (50 µg/mL) was added to the LB medium when required for propagation of the plasmids in *E. coli*. The wild-type (glucose-tolerant) and the recombinant strain of *Synechocystis* 6803 were grown at 30°C in a liquid blue-green medium (BG-11 medium) or on solid BG-11 plates at a light intensity of 100 µmol of photons m⁻²s⁻¹ in ambient air. Kanamycin (20 µg/mL) was added to the BG-11 growth medium as required. Growth of the cells was monitored by measuring their optical density at 730 nm (OD₇₃₀) with an Agilent Cary 60 UV-vis spectrophotometer. 10 mL

cultures for the synthesis of D-lactate were grown (initial OD₇₃₀, 0.4) in 50 mL shake flasks without any antibiotic and 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added for induction. Mixotrophic cultures of *Synechocystis* 6803 were started in BG-11 medium containing a known amount of glucose (0.5%) or acetate (15mM) as an organic carbon source.

5.4.3 Plasmid construction and transformation. The vector pTKA3¹³⁶ served as the backbone for all the plasmids constructed in this study. The gene *gldA101* encoding GlyDH*¹⁹⁰, was amplified from the plasmid pQZ115 with the primers *gldA-o-F2* and *gldA-o-R* (Table 1 and 2). The obtained 1.2 kb fragment was digested with BamHI/NheI and cloned into the same restriction sites of pTKA3, yielding the vector pYY1. A gene cassette, which consists of the codon optimized *gldA101* (*gldA101-syn*) with the promoter P_{trc} in the upstream and the transhydrogenase (*sth*) gene from *Pseudomonas aeruginosa*¹⁹⁸ in the downstream, was chemically synthesized by Genewiz Inc (South Plainfield, NJ) and cloned into the commonly used *E. coli* vector pUC57-kan resulting in the plasmid vector pUC57-glda_sth. The vector pUC57-glda_sth was digested with BamHI/NheI, and the yielding 2.6 kb fragment was cloned into the corresponding restriction sites of pTKA3, resulting in the vector pDY2. The vector pDY3 was constructed by self-ligation of the 8.2 kb fragment obtained through the digestion of pDY2 with KpnI.

Natural transformation of *Synechocystis* 6803 was performed by using a double homologous-recombination procedure as described previously¹⁹. Recombinant colonies appeared between 7 and 10 days post inoculation. The genes of interest were finally integrated into the *psbA1* gene loci (a known neutral site under normal growth conditions) in the genome of *Synechocystis* 6803¹³⁶. For segregation, the positive colonies were propagated continuously

onto BG-11 plates containing kanamycin and segregation of colonies was verified through a colony PCR with the primers AMV17R and ps1_up_fwda (Table 1). The promoter and the heterologous genes in the engineered strains were PCR amplified with respective primers (ptka3-F, CO-F, O-F, sth-F) (Table 1) and sent for sequencing to Genewiz to verify the cloning accuracy.

5.4.4 D (-) lactate analysis. D(-)/L(+) lactic acid detection kit (R-biopharm) was used to measure the D-lactate concentration. Samples of the cyanobacterial culture (50 μ L) were collected every 3 days and centrifuged at 12,000 rpm for 5 min. The supernatant was collected and the D-lactate concentration assay was performed following the manufacturer's instruction. All the reactions were performed in a 96-well plate reader at room temperature (Infinite 200 PRO microplate photometer, TECAN).

5.4.5 ^{13}C isotopomer experiment. To estimate the carbon contributions of glucose and acetate for biomass and D-lactic acid synthesis a ^{13}C labeling experiment was performed. The mutant AV10 was grown in a BG-11 medium with 0.5% glucose (1,2- $^{13}\text{C}_2$ glucose) or 15mM acetate (U- $^{13}\text{C}_2$ acetate) (Sigma, St. Louis). Cultures were started at an OD730 of 0.4 and were grown with labeled glucose or acetate for over 48 hours. The biomass samples and supernatant were collected for measurement of lactate and amino acid labeling.

The proteinogenic amino acids from biomass were hydrolyzed and then derivatized with TBDMS [*N*-(tert-butyldimethylsilyl)-*N*-methyl-trifluoroacetamide], as described previously¹⁶⁰. The derivatized amino acids were analyzed for their ^{13}C mass fraction by GC-MS (Hewlett Packard 7890A and 5975C, Agilent Technologies, USA) equipped with a DB5-MS column

(J&W Scientific) ¹⁶⁰. The fragment [M-57]⁺ containing information of the entire amino acid was used for calculating the ¹³C mass fractions (M: the molecular mass of the derivatized amino acids). The fragment [M-15]⁺ was used only for leucine, since its [M-57]⁺ overlaps with other mass peak ²⁰⁵. To analyze extracellular D-lactic acid labeling, the supernatant (0.2 mL) was first freeze-dried at -50 °C. The dried samples were then pre-derivatized with 200 µL of 2% methoxyamine hydrochloride in pyridine for 60 minutes at 37 °C and then derivatized with 300 µL *N*-Methyl-*N*-(trimethylsilyl) trifluoroacetamide (TMS) for 30 minutes at room temperature. The natural abundance of isotopes, including ¹³C (1.13%), ¹⁸O (0.20%), ²⁹Si (4.70%) and ³⁰Si (3.09%) changes the mass isotopomer spectrum. These changes were corrected using a published algorithm and the detailed measurement protocol can be found in our previous paper ²⁰⁶.

Acknowledgements

We thank Professor K. T. Shanmugam for offering us the plasmid pQZ115. We thank Professor Himadri Pakrasi at WUSTL for his advice on this project. We also thank Diany Li, Kanimozhi, and Zach Hembree for their help with experiments, and Sandra Matteucci from the WUSTL Engineering Communication Center, for her close reading of the manuscript. This research was funded by an NSF Career Grant (MCB0954016).

Authors' contributions

A.M.V conceived the initial idea for this research. A.M.V., Y.Y., and Y.J.T. designed the experiments. A.M.V., Y.Y., and Y.L. performed the experiments. All authors revised and approved the manuscript.

Competing interests

The authors declare competing financial interests since this work is being covered by a pending patent application from Washington University in St. Louis.

Table 5.1: Primer sequences

Primer name	Sequence (5'→3')
gldA-o-F	GGATCCTTGACAATTAATCATCCGGCTCG
gldA-o-F2	GGATCCTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGT GAGCGGATAACAATTTACACAGGAGATATAATCATATGACGAAAA TCATTACCTCTCCAAGCAAGTTTATACAAGG
gldA-o-F3	ATGACGAAAATCATTACCTCTCCAAG
gldA-o-R	GCTAGCTCATGCCATTTTCCTTATAATACCGCCCG
gldA-o-R2	TTAGGCCCACTTTTCCTTGTAATAGC
tranNADH-F	CCTAAGCTAGCGGAGGACTAGCATGG
tranNADH-R	GCTAGCGGTACCTCAAAAAAGCCGG
ptka3-F	CCCGAAGTGCGGAGCCCGAT
CO-F	TTGATGTTGCCTTTGAACCC
O-F	ATGGATACGAAAGTGATTGC
sth-F	GAGCTACCACCTGCGCAACA
AMV17R	GCGCGACTCCCCGTCTTTGACTATCCTTTTTAGGATGGGGCA
ps1_up_fwda	TACCGGAACAGGACCAAGCCTT

Table 5.2: Plasmids and strains

Plasmids/Strains	Description	Source or reference
Plasmids		
pUC57-glda_sth	Chemically synthesized gene cassette consisting of P _{trc} , <i>gldA101-syn</i> and <i>sth</i> .	Genewiz; 190, 198 207
pQZ115	Plasmid carrying <i>gldA101</i>	190
pTKA3	Backbone plasmid for all vectors constructed in this study, with <i>psbA1</i> as the integration loci.	136
pYY1	Derived from pTKA3 with <i>gldA101</i> and the promoter, P _{trc} .	This study
pDY2	Derived from pTKA3 with <i>gldA101-syn</i> , <i>sth</i> and the promoter, P _{trc} .	This study
pDY3	Derived from pTKA3 with <i>gldA101-syn</i> and the promoter, P _{trc} .	This study
Strains		
<i>E. coli</i> 10-Beta	Cloning host strain.	New England Biolabs
<i>Synechocystis</i> sp. PCC 6803	Glucose tolerant wild type, naturally competent.	This study
AV08	<i>Synechocystis</i> P _{trc} :: <i>gldA101</i> ::Km ^r , GlyDH* of <i>Bacillus</i> .	This study
AV10	<i>Synechocystis</i> P _{trc} ::(<i>gldA101-syn</i>)- <i>sth</i> ::Km ^r , GlyDH* of <i>Bacillus</i> , transhydrogenase of <i>Pseudomonas</i> .	This study
AV11	<i>Synechocystis</i> P _{trc} :: <i>gldA101-syn</i> ::Km ^r , GlyDH* of <i>Bacillus</i> .	This study

Figure 5.1A

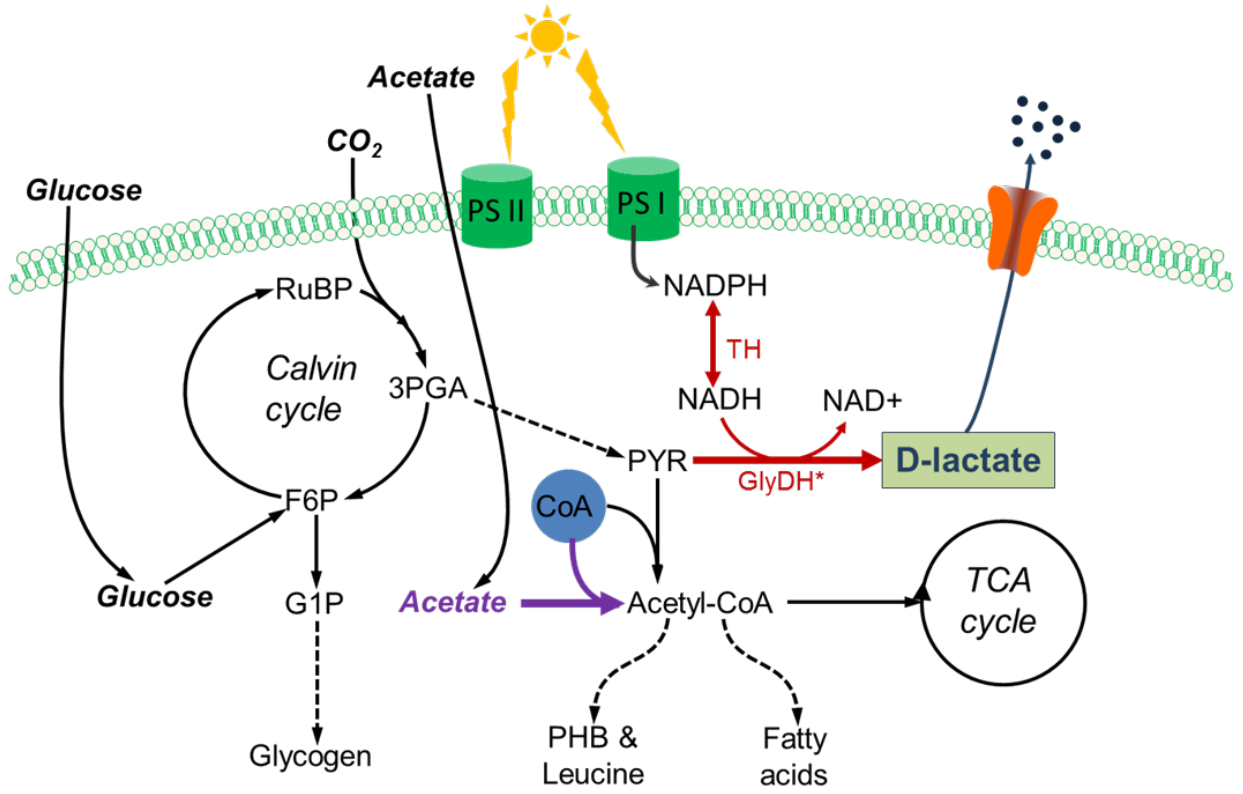


Figure 5.1B

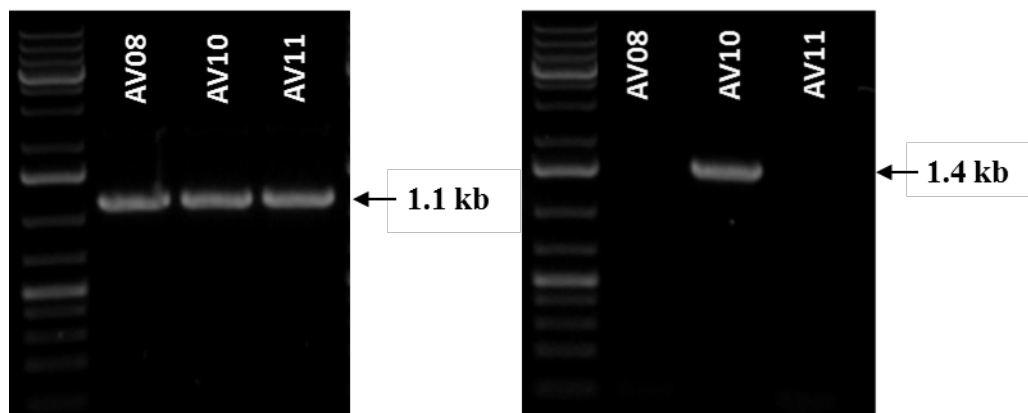


Figure 5.1: Metabolic engineering of *Synechocystis* 6803 for the synthesis of D-lactic acid. (A) Metabolic pathway for D-lactate synthesis. Lactate permeation through the cell membrane occurs either via a lactate transporter or by passive diffusion^{202, 208}. Red arrows indicate the heterologous pathway engineered into *Synechocystis* 6803. **Abbreviations:** GlyDH*, mutant glycerol dehydrogenase; TH, Transhydrogenase; 3PGA, 3-phosphoglycerate; CoA, Coenzyme A; G1P, glucose 1-phosphate; F6P, fructose 6-phosphate; PHB, poly- β -hydroxybutyrate; RuBP, ribulose 1,5-bisphosphate. (B) Colony PCR to verify the presence of the heterologous genes of the mutant glycerol dehydrogenase (Left picture) and transhydrogenase (Right picture) in the engineered strains of *Synechocystis* 6803. *gldA101* was amplified with primers *gldA-o-F3* and *gldA-o-R*; *gldA101-syn* was amplified with primers *gldA-o-F* and *gldA-o-R2*; *sth* was amplified with primers *tranNADH-F* and *tranNADH-R* (Table 5.1).

Figure 5.2A

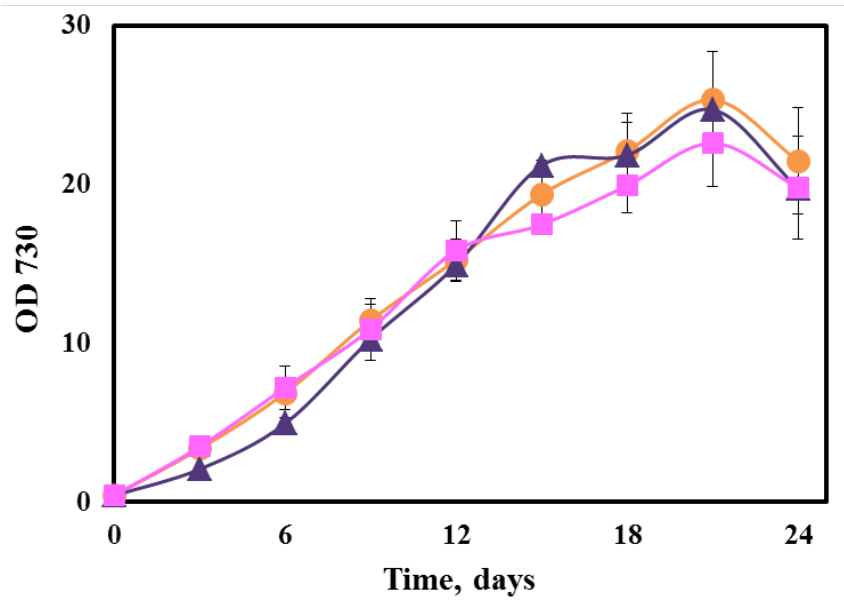


Figure 5.2B

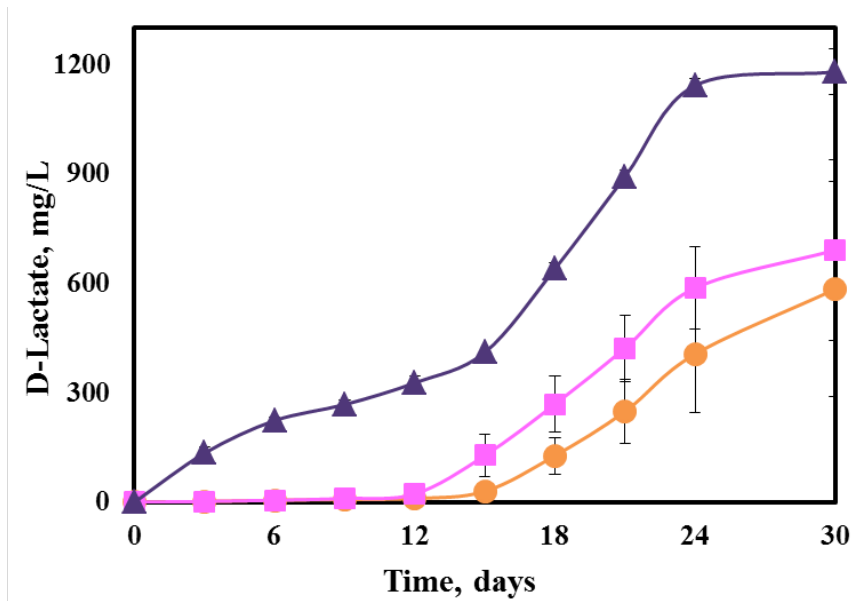


Figure 5.2: Autotrophic production of D-lactate in the engineered strains of *Synechocystis* 6803. (A) Growth curves and (B) D-lactate production in the engineered strains (n = 3). Circles: AV08 (with *gldA101*). Triangles: AV10 (with *gldA101-syn* and *sth*) and Squares: AV11 (with *gldA101-syn*).

Figure 5.3A

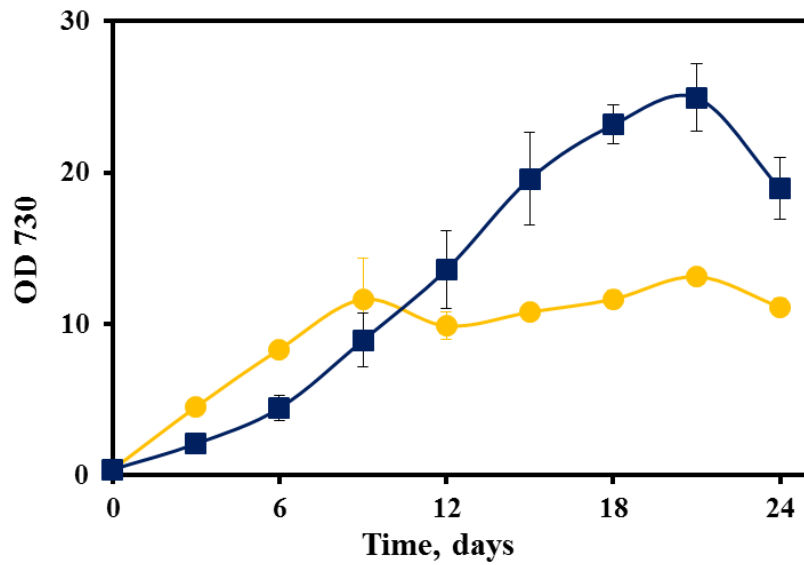


Figure 5.3B

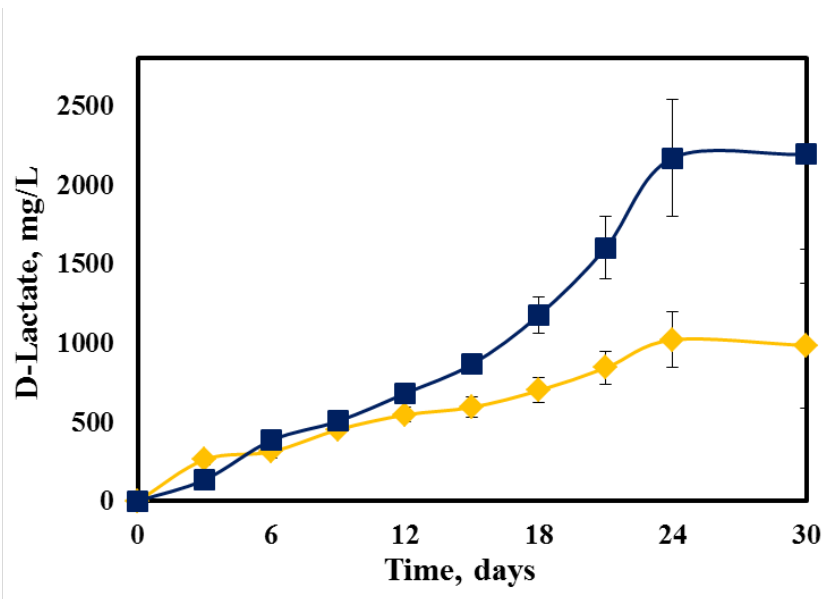


Figure 5.3: Mixotrophic production of D-lactate by AV10. (A) Growth and (B) D-lactate production in the engineered *Synechocystis* 6803 strain AV10 (n = 3), with the provision of additional organic carbon source, i.e., with glucose and acetate (Mixotrophic metabolism). Squares: with acetate. Circles: with glucose.

Figure 5.4A

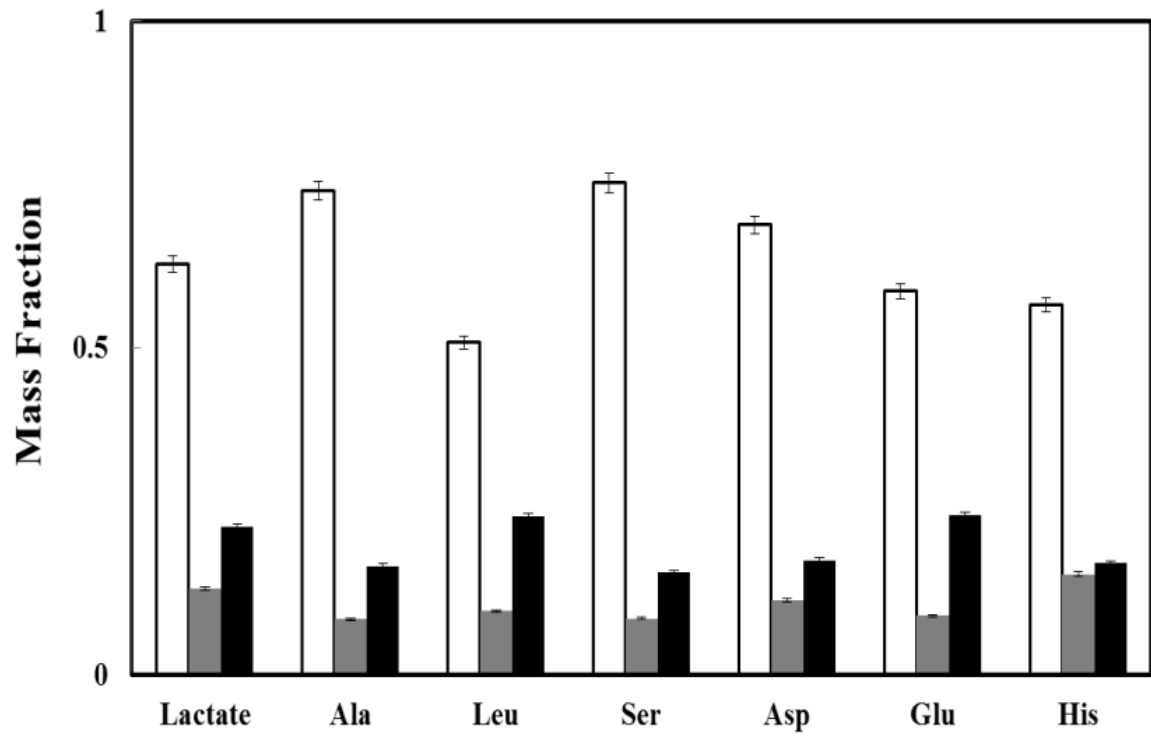


Figure 5.4B

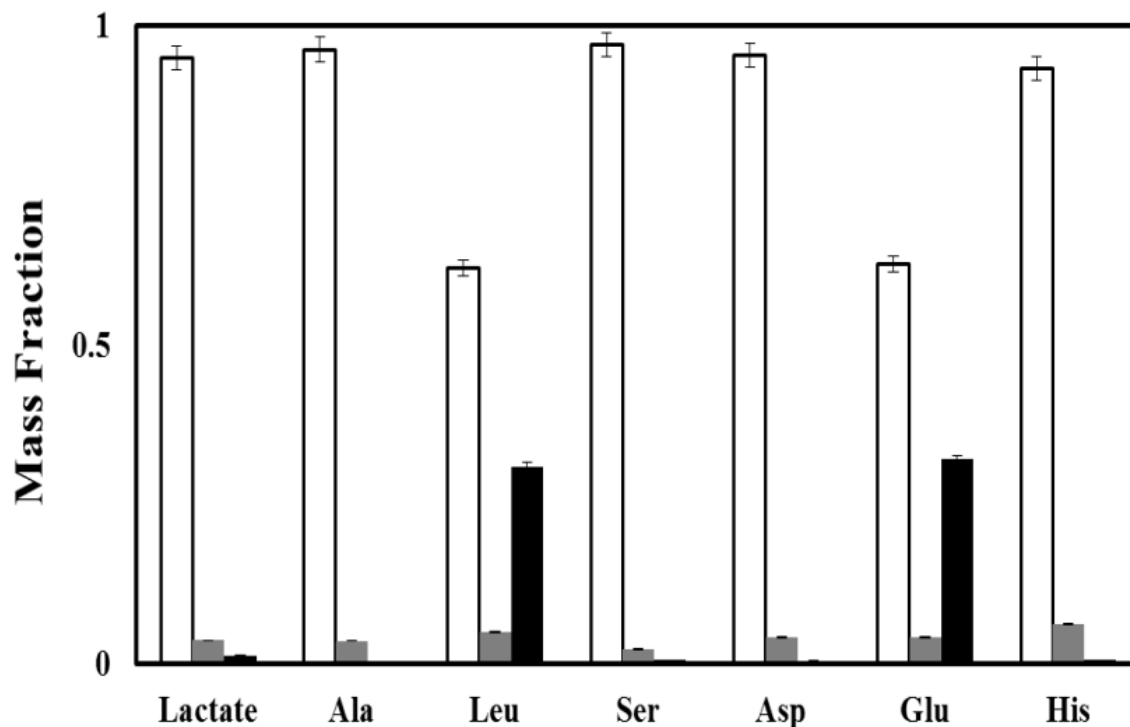


Figure 5.4: Isotopomer analysis showing the mass fraction of isotopomers for selected proteinogenic amino acids [TBDMS based measurement] and D-lactate [MSTFA based measurement]. Standard abbreviations are used for amino acids in the figure. (A) Cultures grown with 5 g/L of [1,2-¹³C] glucose and (B) Cultures grown with 15 mM of [1,2-¹³C] acetate. “white bar” m₀ – mass fraction without any labeled carbon; “grey bar” m₁ – mass fraction with one labeled carbon; “black bar” m₂ – mass fraction with two labeled carbon. (Note: natural ¹³C makes up about 1.1% of total carbon as measurement background)

```

gldA101      1 ATGACGAAAATCATTACCTCTCCAAGCAAGTTTATAACAAGGCCCGATGAATTGTCCAGG
gldA-syn     1 .....C.....C.....C..CTC...A..C..C..G.....C..G.....C..

gldA101     61 CTTTCGGCGTATACGGAAAGGCTTGGCAAAAAAGCATTATTATTGCGGATGATTTTGTGTC
gldA-syn     61 T.GAGT..C.....C...C..T.G.....G..G..C.....C..C..T..C..C..C..G

gldA101     121 ACCGGCCTTGTGCGCAAAACGGTTGAAGAAAGCTATGCCGGCAAAGAAACGGGGTATCAA
gldA-syn     121 .....T.G..G..T..G..T..G..G..GTC.....C..T..C..G

gldA101     181 ATGGCATTATTTCGGTGGTGTAGTGTCTAAACCGGAAATCGAACGGCTTTGTGAAATGAGC
gldA-syn     181 .....C..G..T..C..C..A..CAG...G..T..G.....G..CT.G..C..G...TC.

gldA101     241 AAATCCGAGGAAGCCGATGTCGTTGTGCGAATCGGCGGCGGAAAAACATTGGATACCGCA
gldA-syn     241 .....A.....C..T..G..G..C.....T..C.....C..A..C..T..C

gldA101     301 AAAGCAGTCGGGTATTACAATAACATTCCGGTGATTGTGCGCGCCGACCATCGCTTCCACC
gldA-syn     301 ..G..C..G..C..C.....C.....C..G..C..C..T..T.....

gldA101     361 AATGCCCCGACAAGCGCCCTGTCTGTTATTTACAAAGAGAACGGCGAGTTTGAAGAATAC
gldA-syn     361 .....C..CTC...T..AG...G..C..T..G.....A..C..G.....

gldA101     421 TTGATGCTGCCGCTGAACCCGACTTTTGTTCATTATGGATACGAAAGTGATTGCCTCTGCC
gldA-syn     421 .....T...TT.....T..C.....G..C.....C..T..G.....C.....C..T

gldA101     481 CCTGCCCGCCTGCTCGTTTCCGGCATGGGAGATGCGCTTGCAACGTATTTTGAAGCGCGC
gldA-syn     481 .....T.AT.G.....C..C..CT.G..C..T..C.....C...

gldA101     541 GCCACTAAGCGGGCAAATAAAACGACGATGGCAGGCGGGCGTGTACGGAAGCGGCGATC
gldA-syn     541 .....C.....G..C..C.....T..T..C..C.....T..G..C..C...

gldA101     601 GCGCTTGCAAAACTTTGTTATGACACGCAAATTTGCGAAGGTTTAAAGCAAACTGGCA
gldA-syn     601 ..CT.G..C..GT.G..C..C.....C..G..CAGT..G..C..G..G..C..GT.A..C

gldA101     661 GCGGAAAAACATCTTGTACGGAAGCAGTGGAAAAATCATTGAAGCGAATACGTATCTG
gldA-syn     661 ..C..G.....CT.G..G..C..G..C..T..G..G.....C..G..C..C..T..CT..

gldA101     721 AGCGGAATCGGTTCTGAAAGCGGGCCCTTGCTGCGGCACATGCGATCCATAATGGGCTT
gldA-syn     721 TC...T..T..C..C.....T.....T.A..C..C..T..C..C..T..C..C..CT.G

gldA101     781 ACCGTGCTCGAAGAAACCCATCATATGTACCACGGCGAAAAAGTGGCATTCCGGTACCCTC
gldA-syn     781 .....T.G..G..G.....C..C.....G..G.....T.....C..TT.G

gldA101     841 GCCCAGCTGATTTTGAAGATGCGCCGAAAGCGGAAATTGAAGAGGTGGTCTCCTTCTGC
gldA-syn     841 .....T...C.....G.....C..C..G..C..G.....G.....G.....T

gldA101     901 CTGAGTGTGCGACTTCCCGTACGCTCGGGGATTTGGGCGTGAAAGAACTGAATGAGGAA
gldA-syn     901 T..TCC..G..CT.G..T..T..CT.G..C..C.....G...T...C.....

gldA101     961 AAGCTCCGAAAAGTGGCTGAACTTTCCCTGTGCGGAAGGCGAAACGATTTATAACATGCCG
gldA-syn     961 ..AT.G..G..G.....C...T.G.....C..C..G.....C..C..C.....C

gldA101     1021 TTTGAAGTCACGCCTGACCTTGTGTACGCAATCGTTACCCTGATTCCGTCGGGCGG
gldA-syn     1021 .....G..G..C..C..TT.G.....C..C.....C.....G..C..C

gldA101     1081 TATTATAAGGAAAAATGGGCATGA
gldA-syn     1081 .....C.....G.....C.A.

```

Figure 5.5: Nucleotide sequence alignment of *gldA101* and the codon-optimized *gldA101* (i.e., *gldA101-syn*, synthesized by Genewiz Inc). Conserved nucleotide sequences in *gldA101-syn* are indicated as dotted lines.

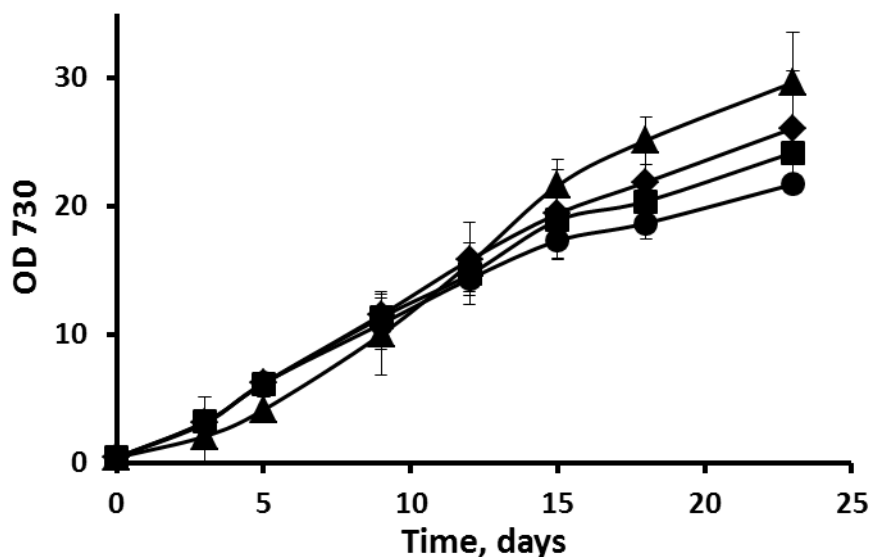


Figure 5.6: Autotrophic growth curve for *Synechocystis* 6803 strains shows similar growth of the engineered D-lactate producing strains as compared to the wild type strain. Diamond: Wild type. Square: AV08. Triangle: AV10. Circle: AV11.

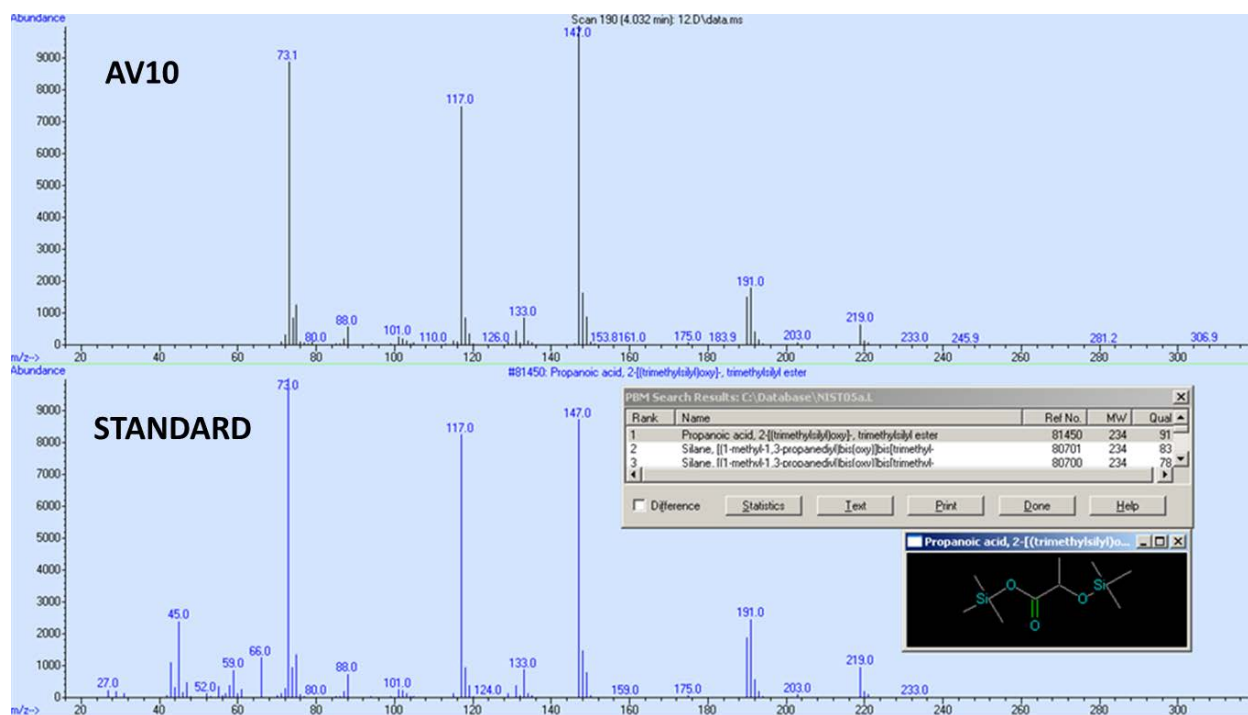


Figure 5.7: Mass spectra obtained via GC-MS confirm the presence of lactate in the cell culture supernatant of AV10 strain. D/L lactate enzyme kit (R-Biopharm) was used to further confirm that the product is an optically pure D-lactate.

Chapter 6: Kinetic modeling and isotopic investigation of isobutanol fermentation by two engineered *Escherichia coli* strains

This chapter has been reproduced from the following publication:

Varman, A.M., Xiao, Y., Feng, X., He, L., Yu, H., & Tang, Y.J. Kinetic Modeling and Isotopic Investigation of Isobutanol Fermentation by Two Engineered *Escherichia coli* Strains. *Industrial & Engineering Chemistry Research* 51, 15855-15863 (2012).

AMV, YX, and XF contributed equally to this work.

Abstract

We constructed an *E. coli* BL21 strain with the Ehrlich pathway (the low performance strain for isobutanol production). We also obtained a high isobutanol-producing *E. coli* strain JCL260 from the James Liao group (University of California). To compare the fermentation performances of the two engineered strains, we employed a general Monod-based model coupled with mixed-growth-associated isobutanol formation kinetics to simulate glucose consumption, biomass growth, and product secretion/loss under different cultivation conditions. Based on both kinetic data and additional ¹³C-isotopic investigation, we found that the low performance strain demonstrated robust biomass growth in the minimal growth medium (20 g/L glucose), achieving isobutanol production (up to 0.95 g/L). It utilized significant amount of yeast extract to synthesize isobutanol when it grew in the rich medium. The rich medium also enhanced waste product secretion, and thus reduced the glucose-based isobutanol yield. In contrast, JCL260 had poor biomass growth in the minimal medium due to an inflated Monod constant (K_s), while the rich medium greatly promoted both biomass growth and isobutanol productivity. With the optimized keto acid pathway, JCL260 synthesized isobutanol mostly from glucose even in the

presence of sufficient yeast extract. This study not only provided a kinetic model for scaled-up isobutanol fermentation, but also offered metabolic insights into the performance tradeoff between two engineered *E. coli* strains.

Key words: ¹³C-isotopic, Ehrlich pathway, mixed-growth-associated, tradeoff, yeast extract

6.1 Introduction

Biobutanols are second generation biofuels that have higher energy density and lower water solubility than ethanol. Acetone-butanol-ethanol (ABE) fermentation is a traditional bioprocess that uses *Clostridium acetobutylicum* to produce *n*-butanol, but such a process is restrained by the slow alcohol production rate²⁰⁹. To overcome this restriction, the *n*-butanol pathway derived from *Clostridium* has been reconstructed in fast-growing *E. coli* or yeast strains^{23, 24, 118, 210}. Butanol biosynthesis via the *Clostridium* pathway has limitations including low product titer and yield due to the accumulation of toxic metabolites. Another approach is via the keto-acid pathway to produce low-toxicity isobutanol (IB),²⁵ where the amino acid biosynthesis pathways and the Ehrlich pathway are incorporated for alcohol synthesis^{26, 27}. This method shows effective production of higher alcohols because of robust and ubiquitous amino acid pathways.

Table 6.1 summarizes diverse biobutanol production strategies, including the overexpression of the targeted pathway in different microbial hosts (including photoautotrophic microbes), the elimination of competing pathways, the systems redesign of host metabolism, and the integration of fermentation with *in situ* product separation. However, few papers have studied the kinetics of engineered microbial hosts for biobutanol fermentation. To apply a newly developed host in the biofuel industry, a kinetic-based model is of practical importance not only

for designing optimal scaled-up fermentation, but also for understanding the internal metabolic features of microbial hosts in responses to various nutrient sources and cultivation conditions. To fulfill this gap, our lab has created an *E. coli* mutant that produces IB via the Ehrlich pathway. Meanwhile, we have obtained a high performance *E. coli* strain JCL260 with an optimized metabolism for IB synthesis (offered by the James Liao group)¹³⁰. Based on fermentation data using both strains, we developed an empirical model to analyze and compare their fermentation kinetics. We also performed ¹³C-experiments to investigate the nutrient use of the two mutant strains for the synthesis of biomass and IB.

6.2 Materials and Methods

6.2.1 Pathway construction

We engineered *E. coli* BL21 (DE3) by heterologous expression of Kivd (2-ketoisovalerate decarboxylase) and AdhA (aldehyde reductase). The two genes were amplified from *Lactococcus lactis* by PCR with high fidelity DNA polymerase Pfx (Invitrogen). Primers for *kivd*: 5'-gacactcgagtaatgtatacagtaggagattac-3'; 5'-tgcggttaccttatgattttttgttc-3'. Primers for *adhA*: 5'-tcaactagtggaccaggagatataatatgaaagcagcagtagtaagac-3'; 5'-atttgcggccgcgcatgcttatttagtaaaatcaatgac-3'. The genes *kivd* (treated with XhoI / KpnI) and *adhA* (treated with KpnI / SphI) were cloned into the pTAC-MAT-Tag-2 Expression Vector (Sigma-Aldrich) via XhoI / SphI to create the plasmid pTAC-KA, and then transformed into *E. coli* BL21 (DE3). This low performance mutant used its native valine biosynthesis pathway to generate 2-ketoisovalerate, and then converted it to IB by the heterologous Ehrlich pathway (Fig. 6.1). To confirm the expression of Kivd and AdhA, we performed SDS-PAGE analysis of the recombinant strain and observed the protein bands of Kivd (~ 60 kDa) and AdhA (~ 35 kDa).

The strain secreted IB, acetate, lactate, ethanol, and a small amount of n-propanol and methylbutanol (similar product profiles to other IB producing *E. coli* strains)²⁵. Additionally, Professor James Liao from University of California offered us an *E. coli* strain JCL260 with plasmids pSA65 and pSA69²¹¹. This high performance strain not only contains two plasmids that overexpress the entire IB pathway, but also has gene deletions to interfere with waste product (acetate, formate, ethanol, succinate, and lactate) biosynthesis.

6.2.2 Fermentation conditions

Fermentations were performed in a New Brunswick Bioflo 110 fermentor with a dissolved oxygen (DO) electrode, a temperature electrode, and a pH meter. The 100% DO was defined as the point where the cell-free medium was purged by air (~2 L/min) for 15 minutes. In the oxygen limited fermentations (air rate = 0 L/min), the DO dropped to 0% during the exponential growth phase. Two culture media were used: (a) a minimal medium that contained 2% glucose, M9 salts (Difco), 10 mg/L vitamin B1, and 50 mg/L ampicillin; and (b) a rich medium containing the minimal medium with 5 g/L yeast extract. To start each fermentation, 400 ml of culture was inoculated with 5 ml of overnight LB culture (OD₆₀₀~3) of the recombinant *E. coli* strain. The cultivation conditions were: pH = 7.0 (controlled by adding 2 mol/L NaOH via an auto-pump), temperature = 30 °C, and stirring speed = 200 rpm. For all fermentations, cells were first grown in aerobic conditions (DO>50%) before adding 0.2 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside). Right after IPTG induction, we imposed two O₂ conditions: 1) in aerobic conditions, air (flow rate: ~1 L/min) was bubbled into the bioreactor to provide O₂ and to remove IB (i.e., gas stripping) from the bioreactor; 2) in O₂ limited conditions, air was turned off and the DO was maintained zero during IB production. For the low performance strain, we had three fermentations: F1 (minimal medium and aerobic conditions), F2 (minimal medium and O₂

limited conditions), and F3 (rich medium and O₂ limited conditions). For JCL260, we had two fermentations (F4: minimal medium and aerobic conditions; F5: rich medium and aerobic conditions).

6.2.3 Analytical methods for biomass and metabolites

Culture samples were taken after IPTG induction, ~3 ml of culture was taken from the bioreactor at each time point for metabolite and biomass analysis. Biomass growth was monitored by optical density OD₆₀₀. There was a linear relationship between the dry cell weight and OD₆₀₀. To measure dry biomass weight, biomass samples were harvested by centrifugation, washed with DI water, and dried at 100°C until their weight remained constant. Glucose, ethanol, acetate, and lactate were measured using enzyme kits (R-Biopharm). Alcohols could be detected using GC (Hewlett Packard model 7890A, Agilent Technologies, equipped with a DB5-MS column, J&W Scientific) and a mass spectrometer (5975C, Agilent Technologies). The GC-MS detected ethanol, IB, propanol and methyl-butanol. The IB concentration was determined by a modified GC-MS method.¹¹⁸ Briefly, 400 µl of supernatant was extracted with 400 µl of toluene (Sigma-Aldrich) by 2-min vortex, followed by high-speed centrifugation (16000×g). The organic layer was taken for GC-MS analysis under the following program: hold at 70 °C for 2 min, ramp to 230 °C at 20 °C min⁻¹, and then hold at 300 °C for 6 min. The carrier gas was helium. The MS scan mode was from m/z 20 to 200. Samples were quantified relative to a standard curve of IB concentrations for MS detection, and methanol was taken as an internal standard.

6.2.4 ¹³C-experiments for analyzing nutrient contributions to isobutanol productions

In the ¹³C-experiments, the minimal medium with 2% fully labeled glucose (Cambridge Isotope Laboratories) was supplemented with 1 g/L or 5 g/L yeast extract (Bacto). By measuring

^{13}C -abundance in key metabolites from the engineered strains, we estimated the contribution of yeast extract (non-labeled) to biomass and IB synthesis in the ^{13}C -glucose medium. Specifically, 5 ml of cultures (with ^{13}C -glucose and yeast extract) were inoculated with 5 μl of overnight LB culture of the engineered strain in a 50 ml falcon tube with a closed cap (shaking at 200 rpm, 30°C). The cultures (JCL260 or the low performance strain) were induced by 0.2 mM IPTG (when $\text{OD}_{600} > 0.2$), and the samples were taken (at $t \sim 24$ hours, middle-log growth phase) for isotopomer analysis of IB and amino acids. The two mass-to-charge peaks ($m/z=74$ for unlabeled IB and $m/z=78$ for labeled IB) were quantified. Their ratio approximately corresponded to the ratio of IB synthesized from unlabeled yeast extract vs. labeled glucose. Concurrently, we did isotopic analysis of proteinogenic amino acids to identify the incorporation of unlabeled carbon from yeast extract into biomass protein. The measurements were based on a GC-MS protocol, using TBDMS (N-(tert-butyldimethylsilyl)-N-methyl-trifluoroacetamide, Sigma-Aldrich) to derivatize hydrolyzed amino acids from the biomass ¹⁶¹. The m/z ions $[\text{M}-57]^+$ from unfragmented amino acids were used for analysis except leucine and isoleucine. Because of overlapping ions with $[\text{M}-57]^+$, the $[\text{M}-159]^+$ was used to calculate the isotopomer labeling information of leucine and isoleucine ¹⁶².

6.2.5 Model formulation

We developed a kinetic model to describe the fermentation data after IPTG induction. The model contained six time-dependent process variables: X, ACT, LACT, EtOH, IB, and Glu, which represented the concentrations of biomass, acetate, lactate, ethanol, IB, and glucose, respectively. The biomass growth model consisted of glucose-associated (R_X) and yeast-extract-associated ($R_{X,YE}$) terms. IB production was simulated by a mixed-growth-associated product formation model (Eq. 6.5), where β was the non-growth associated IB production rate. In Eq.

6.1~6.6, k_d was the cell death rate; Y_{AL} was the acetate yield from lactate (equal to 0.67 g ACT/g LACT, based on a 1:1 mol ratio); Y_{XG} , Y_{AG} , Y_{EG} , Y_{LG} , and Y_{IG} were the growth associated glucose yields to biomass, acetate, ethanol, lactate, and IB. k_{IB} was the removal rate of IB due to gas stripping under aerobic fermentation F1, F4 and F5. In F2 and F3, IB loss was minimal (k_{IB} was set to zero). A first-order kinetic parameter (k_{act}) was used to describe acetate production from lactate.

$$\frac{dX}{dt} = R_X - k_d \cdot X + R_{X,YE} \quad (6.1)$$

$$\frac{dACT}{dt} = R_A + Y_{AL} \cdot k_{act} \cdot LACT \cdot X \quad (6.2)$$

$$\frac{dLACT}{dt} = R_L - k_{act} \cdot LACT \cdot X \quad (6.3)$$

$$\frac{dEtOH}{dt} = R_E \quad (6.4)$$

$$\frac{dIB}{dt} = R_{IB} + \beta \cdot X - k_{IB} \cdot IB \quad (6.5)$$

$$\frac{dGlu}{dt} = -\frac{R_X}{Y_{XG}} - \frac{R_A}{Y_{AG}} - \frac{R_E}{Y_{EG}} - \frac{R_L}{Y_{LG}} - \frac{R_{IB}}{Y_{IBG}} \quad (6.6)$$

In Eq. 6.7~ 6.12, R_X , R_A , R_E , R_L , and R_{IB} were the production rates of biomass, acetate, ethanol, lactate, and IB from glucose, respectively.

$$R_X = \frac{\mu_{max,app} \cdot Glu}{K_S + Glu} \cdot \frac{1}{1 + \frac{ACT}{K_{iA}}} \cdot X \quad (6.7)$$

$$R_A = \alpha_{AX} \cdot R_X \quad (6.8)$$

$$R_E = \alpha_{EX} \cdot R_X \quad (6.9)$$

$$R_L = \alpha_{LX} \cdot R_X \quad (6.10)$$

$$R_{IB} = \alpha_{IBX} \cdot R_X \quad (6.11)$$

$$R_{X, YE} = \mu_{max,YE} \cdot e^{-k_{YE} \cdot t} X \quad (6.12)$$

R_x represented a growth model with Monod constant K_S and maximum specific rate coefficient $\mu_{\max,app}$. Since acetate inhibited *E. coli* growth by decreasing the intracellular pH, a non-competitive inhibition K_{iA} was included in the model²¹². The dependence of the glucose-based growth rate on oxygen (i.e., aerobic growth vs. anaerobic growth) was implicitly included in the calculation of $\mu_{\max,app}$ (i.e., the oxygen conditions affected $\mu_{\max,app}$). α_{AX} , α_{EX} , α_{LX} , and α_{IBX} were the growth-associated yields of acetate, ethanol, lactate, and IB, respectively. In the rich medium, the yeast extract was quickly consumed to support biomass growth. The model included a yeast-extract-associated biomass growth rate $R_{X,YE}$ using a two-parameter exponential decay function Eq. 6.12. Table 6.2 summarized model parameters and their units.

For each batch culture, unknown parameters were determined by minimizing the sum of the squares of the differences between the model's predictions and the experimentally observed growth and metabolite profiles²¹³. The "ode23" command in MATLAB (R2009a, Mathworks) solved the differential equations, while the "fmincon" command searched suitable values of parameters. To reduce the risk of having local solutions during the nonlinear parameter estimation, we tested the initial guesses for 30 times within the range of possible values to identify the global solution. To evaluate the quality of the parameter estimates, we checked the sensitivity of the estimated parameters to the measurement inaccuracies. Fifty simulated fermentation data sets (including both biomass and metabolite data) were generated by the addition of normally distributed measurement noise to the fermentation data set (i.e., randomly perturbed the measured data by 30%). The same data-fitting algorithm found new sets of parameters. From the probability distribution of these parameter distributions, standard deviations of model-fitted parameters were estimated.

6.3 Results and discussion

6.3.1 Isobutanol fermentation results

In this study, both engineered *E. coli* strains employed the Ehrlich pathway (Fig. 6.1), where 2-ketoisovalerate from valine metabolism is redirected to IB synthesis. For the low performance strain, we simply over-expressed 2-ketoisovalerate decarboxylase and alcohol dehydrogenase. For strain JCL260, both the Ehrlich pathway and 2-ketoisovalerate synthesis pathway were overexpressed. This strain also had gene deletions involved in by-product formation to increase pyruvate for IB synthesis, so it was reported to produce 22 g/L of IB in 112 hrs.²⁵

This study compared IB fermentation kinetics between the two strains. For the low performance strain, ethanol and lactate were barely detected in the aerobic conditions (Fig. 6.2). IB titer only reached (0.2 g/L) in F1, because the in situ removal of IB was considerable (the airflow carried IB out of the fermentor). Such gas stripping is an effective strategy to avoid the IB accumulation in the culture that causes the inhibitory effect on alcohol production²¹¹. In O₂ limited conditions, the F2 generated 0.95 g/L IB, 1.5 g/L ethanol, 2.2 g/L acetate, and 5.1 g/L lactate, while the lactate was reused in the late fermentation stage (stationary growth phase). With the addition of yeast extract, the F3 had fast biomass growth (Fig. 6.4). The cell density reached a peak (2 g DCW/L biomass) after seven hours of IPTG induction, and glucose was consumed within ~12 hours (compared to ~40 hours in the F1 and F2). The high rates for biomass growth promoted IB production rate. It took the F3 15 hours to generate 0.6 g/L IB, while it took F2 40 hours to generate same amount of IB. The addition of yeast extract also resulted in a large amount of growth-associated organic acids (6.0 g/L lactate and 3.6 g/L

acetate), and thus decreased IB yield from glucose (0.7 g/L IB and 2.0 g/L ethanol from the F3). A recent paper reported that JCL260 accumulated up to 7 g/L IB in an aerobic batch culture using the culture media containing 55 g/L glucose, 2.2 g/L sodium citrate, 25 g/L yeast extract, and complex trace metal solution ²¹¹. This study performed two aerobic fermentations using JCL260. In the complete minimal medium with 20 g/L glucose (F4, Fig. 6.5), JCL260 had very slow biomass growth and low IB titer (~0.1 g/L). When yeast extract (5 g/L) was supplemented (F5, Fig 6.6), IB productivity was significantly improved and its titer reached ~1 g/L (over fivefold higher than the low performance strain). Meanwhile, JCL260 produced only 1 g/L acetate (two times lower than the low performance strain) because of the deletion of phosphotransacetylase (*pta*) ²¹¹.

6.3.2 Kinetic modeling of isobutanol fermentation

The same kinetic model simulated fermentation processes by two IB producers. Table 6.2 lists the kinetic parameters obtained by nonlinear parameter fitting. For the low performance strain, the specific growth rate $\mu_{\max,app}$ (0.015 h⁻¹) in the oxygen limited conditions was lower than that in the aerobic culture conditions (0.051 h⁻¹). IB could be synthesized in both growth and stationary phases. The O₂ limited condition reduced growth associated IB yield, but promoted non-growth associated IB production (e.g., $\beta = 0.012$ g IB/g biomass·h in the F2). In the presence of yeast extract, the yeast extract associated biomass growth rate ($\mu_{\max,YE}=0.48$ hr⁻¹) was one order of magnitude higher than glucose-associated growth rates. The addition of yeast extract (F3) also improved the biomass yield coefficient ($Y_{XG} = 0.20$) and the growth associated IB production ($\alpha_{IBX} = 0.78$ g IB/g biomass). Meanwhile, the yeast extract increased yield coefficients of waste products (Y_{AG} , Y_{EG} , Y_{LG}) in the F3. The IB yield coefficient Y_{IG} was 0.26 g IB/g glucose under aerobic respiration, higher than Y_{IG} under O₂ limited conditions (F2 and F3).

For JCL260, the fermentation data indicated that the strain had a highly inflated Monod constant K_s (10 g/L), which caused the biomass growth rate to be slower than that of the low performance strain. The slow growth led to poor IB synthesis in F4 ($\alpha_{IBX}=0.06$ g IB/g biomass). Because of the knockout of the *pta* gene to reduce acetate synthesis, the growth associated acetate production α_{AX} in F4 was 0.35 g acetate/g biomass, suggesting that acetate production rate was reduced compared to the low performance strain ($\alpha_{AX}=0.62$ g acetate/g biomass in the F1). On the other hand, JCL260 still generated acetate after *pta* deletion²¹¹. The alternate acetate pathways in JCL260 had higher glucose associated acetate yield (Y_{AG}) than that of the low performance strain under aerobic conditions. This observation was consistent with the fact that JCL260 (the strain with multiple gene knockouts) had a poor respiration rate, and thus higher fraction of glucose was converted to biomass (i.e., Y_{XG} also increased) and byproducts rather than degraded to CO_2 . When yeast extract was added to the growth media, the growth associated IB production α_{IBX} was 3.3 g IB/g biomass, which was about 5.7 folds higher than that of the low performance strain. The addition of nutrients improved the JCL260 biomass growth, the cell energy (such as NADH) generation, and the carbon flux through the IB pathway. In contrast, the low performance strain had a suboptimal IB pathway. Therefore, yeast extract only enhanced metabolic overflows to waste metabolites rather than improving IB titers (the F3).

Finally, the continuous flow of air into the bioreactor performed an in situ stripping of IB out of the reactor in the aerobic conditions (F1, F4 and F5). Using the model, we estimated the total IB production by JCL260 without any loss by gas stripping (i.e., $k_{IB} = 0$, Fig. 6.6). The model showed that the total IB could reach 5 g/L in F5. This result indicated that the IB production can be significantly improved via the integration of IB fermentation with a downstream product recovery process.

6.3.3 Analysis of the role of yeast extract for isobutanol synthesis

Nutrient supplements play an important role in improving fermentation performance. Rich media have been commonly used for butanol fermentations^{25, 118, 214}. In addition to providing the building blocks for biomass growth, *E. coli* can also utilize the Ehrlich pathway to convert protein hydrolysates to higher alcohols²¹⁵. However, the contribution of rich nutrient (yeast extract) to IB production was not quantified. Here, we used ¹³C-experiments to determine the ratio of carbon utilization from two different sources (nonlabeled yeast extract vs. fully labeled ¹³C-glucose) under oxygen limited conditions via GC-MS analysis (Fig. 6.7). For the low performance strain cultivation with 1 g/L yeast extract, its proteinogenic amino acids (e.g., histidine, leucine, isoleucine, lysine, and proline) were highly imported from exogenous amino acids (>50%, corresponding to the ¹²C-dilutions), while IB was mostly labeled with four carbons (m/z=78, IB came from labeled glucose). When excess yeast extract (5 g/L) was provided, the low performance strain not only used yeast extract as the building blocks for cell growth, but also converted it to IB (~50% IB was nonlabeled). On the other hand, with sufficient yeast extract (5 g/L), JCL260 still mainly used ¹³C-glucose for IB production (labeled IB was > 90%). In the rich media, JCL260 highly utilized yeast extract for biomass synthesis. It showed much higher ¹³C-labeling concentration (~20%) in valine than the low performance strain (~5%). Higher abundance of ¹³C-labeling in valine proved that the overexpression of the keto acid pathway in JCL260 efficiently enhanced the ¹³C-glucose flux towards 2-ketoisovalerate (the common precursor for both IB and valine) and reduced the relative valine uptake from the rich media.

6.4 Concluding remarks

This study developed a general empirical model for IB fermentations by two engineered *E. coli* strains. The model with nonlinear fitted parameters reasonably well described batch fermentation data under denoted cultivation conditions. The model results indicated that the two strains displayed a difference in biomass growth behavior and products generation. The comparative study revealed the change of influential kinetic variables in responses to the cultivation conditions. Moreover, we quantified the contribution of nutrient sources to product yields via isotopic investigation, and proved that the keto-acid pathway was a rate-limiting step for IB production in the low performance strain. This study may serve as a springboard for developing useful bioprocess models for higher alcohols fermentations in the biotechnology industry.

Abbreviations

Abbreviations: ACA, acetyl-CoA; AdhA, aldehyde reductase; ALA, 2-acetolactate; DHI, 2,3-dihydroxy-isovalerate; G3P, glyceraldehyde 3-phosphate; G6P, glucose 6-phosphate; Glu, glucose; IB, isobutanol; IBA, isobutanal; KIV, 2-keto-isovalerate; Kivd, 2-keto-isovalerate decarboxylase; PYR, pyruvate; Val, valine; PP pathway, pentose phosphate pathway; TCA cycle, tricarboxylic acid cycle.

Acknowledgement

This study was supported by an NSF Career Grant (MCB0954016) and the Clean Coal Consortium at Washington University. We thank James Liao for offering us the JCL 260 strain

for this study. We also thank Chunlei Mei for her technical assistance. The modeling of the low performance strain was done by XF, while the modeling of the JCL260 was done by AMV. The authors have declared no conflict of interest.

Table 6.1: Recent studies on biobutanol production by engineered microorganisms

Products	Substrate	Host cell	Titer	Research Highlights	Ref
IB	Glucose	<i>E. coli</i>	22 g/L	Introduction of a non-fermentative pathway to produce IB; elimination of competing pathways to reduce waste metabolite secretion	25
IB	Glucose	<i>E. coli</i>	50 g/L	<i>In situ</i> IB removal from the bioreactor using gas stripping	211
IB	CO ₂	<i>Synechococcus elongatus</i>	~0.4 g/L	Overexpression of both non-fermentative pathway and <i>Rubisco</i> for autotrophic IB production	28
IB	Cellulose	<i>Clostridium cellulolyticum</i>	0.66 g/L	Direct conversion of cellulose to IB using engineered cellulolytic bacterial species	10
IB	Glucose	<i>E. coli</i>	1.7 g/L	A strain optimized for IB production via elementary mode analysis	216
IB	Glucose	<i>E. coli</i>	13.4 g/L	Utilization of the NADH-dependent enzyme in keto-acid pathway to alleviate co-factor imbalance	114
IB	Amino acids	<i>E. coli</i>	~2 g/L	Utilization of protein hydrolysates for higher alcohols synthesis by introducing enzymes for exogenous transamination and deamination cycles	215
IB	CO ₂	<i>Ralstonia eutropha</i>	~1 g/L	Developing an electromicrobial process to convert CO ₂ to higher alcohols	217
Butanol	Glucose	<i>E. coli</i>	1 g/L	A strain engineered for 1-butanol and 1-propanol production via isoleucine biosynthesis pathway	218
Butanol	Galactose	<i>Saccharomyces cerevisiae</i>	2.5 mg/L	Overexpression of <i>n</i> -butanol pathway derived from <i>Clostridium</i>	
Butanol	Glucose	<i>E. coli</i>	4.6 g/L	Increase of the barrier for the reverse reaction of butyryl-CoA to crotonyl-CoA via trans-enoyl-CoA reductase	118
Butanol	Glucose	<i>E. coli</i>	30 g/L	Use of trans-enoyl-CoA reductase and optimization of NADH & acetyl-CoA driving forces	24
Butanol	CO ₂	<i>Synechococcus elongatus</i>	14.5 mg/L	Anaerobic production of 1-butanol from CO ₂ using CoA-dependent butanol pathway	29
Butanol	Glucose	<i>E. coli</i>	~14 g/L	Utilization of a functional reversal of the beta-oxidation cycle for the synthesis of alcohols	109
Butanol	CO ₂	<i>Synechococcus elongatus</i>	30 mg/L	Driving butanol synthesis pathway forward via an engineered ATP consumption	175

Table 6.2: Parameters of Monod model for *E. coli* IB fermentation

	Notations	Units	F1	F2	F3	F4	F5
K_S	Monod constant	g/L	0.32±0.05 ^a	0.32±0.05	0.32±0.05	10 ±1 ^c	10 ±1
K_{iA}	Acetate inhibition	g/L	49±11 ^a	49±11	49±11	1.0 ± 0.2 ^c	1.0 ± 0.2
$\mu_{max,app}$	Specific growth rate	/h	0.051±0.004	0.015±0.001 ^b	0.015±0.001	0.12± 0.01 ^c	0.12± 0.01
Y_{XG}	Biomass yield from glu	g biomass/g glu	0.18±0.03	0.14±0.01	0.20±0.04	0.38±0.01	0.39±0.03
Y_{AG}	Acetate yield from glu	g acetate/g glu	0.076±0.007	0.083±0.004	0.33±0.07	0.32±0.01	0.35±0.02
Y_{EG}	Ethanol yield from glu	g ethanol/g glu	NA	0.26±0.01	0.40±0.05	NA	NA
Y_{LG}	Lactate yield from glu	g lactate/g glu	NA	0.56±0.01	0.91±0.10	NA	NA
Y_{IG}	IB yield from glu	g IB/g glu	0.26±0.05	0.033±0.001	0.19±0.04	0.22±0.01	0.36±0.02
α_{AX}	Growth associated acetate synthesis	g acetate/g biomass	0.62±0.02	0.30±0.01	3.0±0.2	0.35±0.01	0.51±0.03
α_{EX}	Growth associated ethanol synthesis	g ethanol/g biomass	NA	3.7±0.2	4.0±0.2	NA	NA
α_{LX}	Growth associated lactate synthesis	g lactate/g biomass	NA	14±1	14±1	NA	NA
α_{IBX}	Growth associated IB synthesis	g IB/g biomass	0.58±0.05	0.078±0.01	0.78±0.06	0.06±0.01	3.3±0.1
k_d	Cell death rate	/h	0.010±0.002	0.001±0.0002	0.010±0.001	0.02±0.01	0 ± 0.001
k_{IB}	gas stripping rate	/h	0.11±0.02	NA	NA	0.11±0.01	0.11±0.01
k_{act}	acetate production from lactate	(h·g biomass/L) ⁻¹	NA	0.013±0.001	0.0034±0.0002	NA	NA
k_{YE}	Yeast extract consumption rate	/h	NA	NA	0.55±0.03	NA	0.65±0.05
$\mu_{max,YE}$	Apparent specific growth rate with yeast extract	/h	NA	NA	0.48±0.03	NA	0.32±0.03
β	Non-growth associated IB production	g IB/(g biomass·h)	0.002±0.002	0.012±0.001	0.006±0.0	0±0.0	0±0.0

a): model assuming same values for F1, F2, and F3. b): model assuming same values for F2 and F3. c): model assuming same values for F4 and the F5. NA: not applicable.

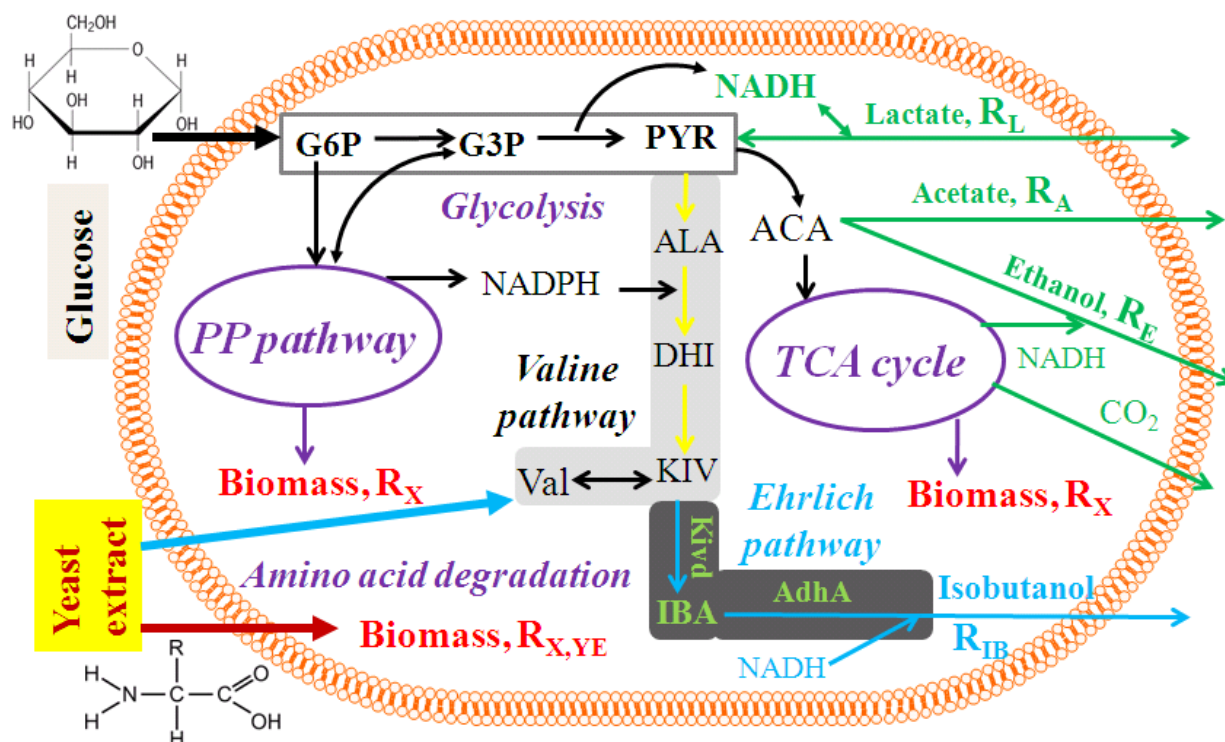


Figure 6.1: Metabolism in the *E. coli* strains for IB production. R_X , $R_{X,YE}$, R_A , R_E , R_L , and R_{IB} were shown in the Equations 6.1~6.12. IB synthesis consumes one mole NADPH (by keto-acid reductoisomerase) and one mole NADH (by aldehyde reductase). The cell met metabolism removes the redundant NADH by O_2 oxidation or by synthesis of lactate and ethanol.

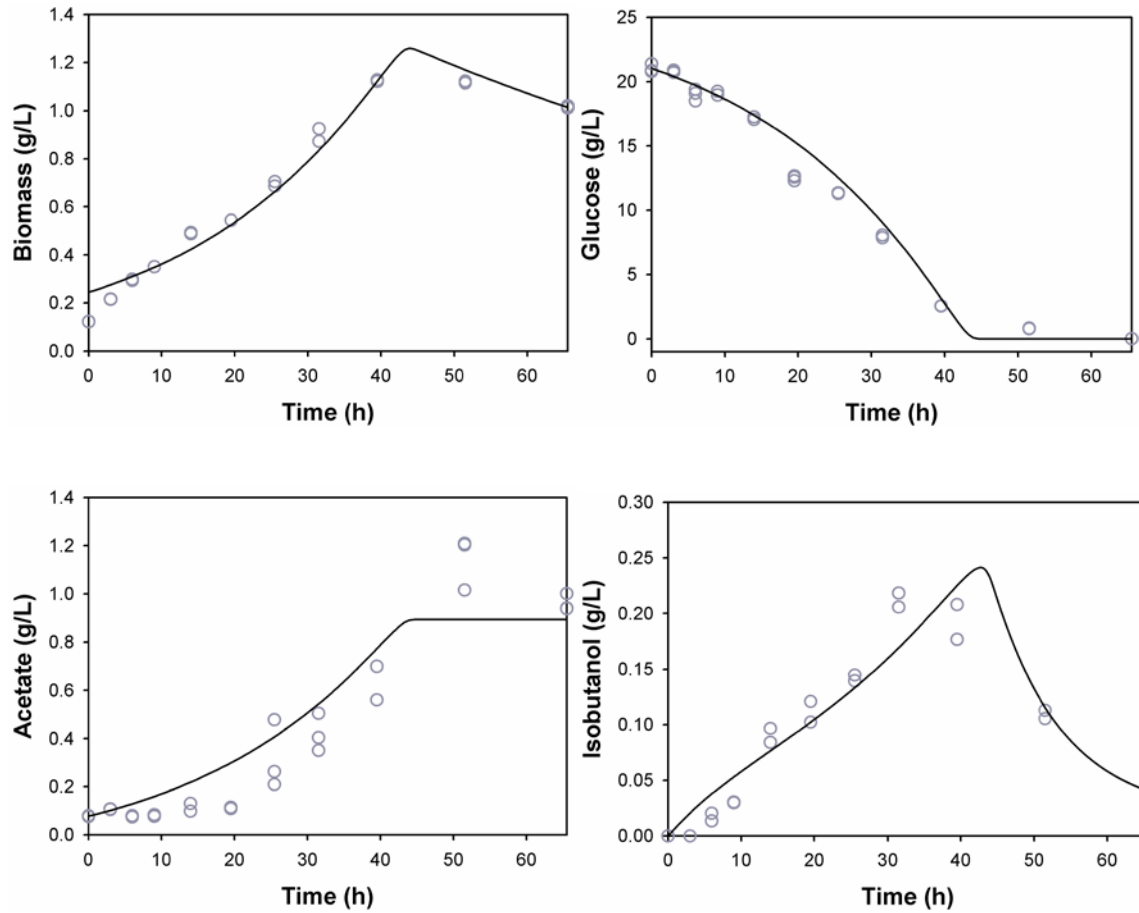


Figure 6.2: Growth kinetics after IPTG induction (F1). The circles were experimental measurements, and the solid lines were simulations from the Monod kinetic model (same as Fig. 6.3~Fig. 6.6).

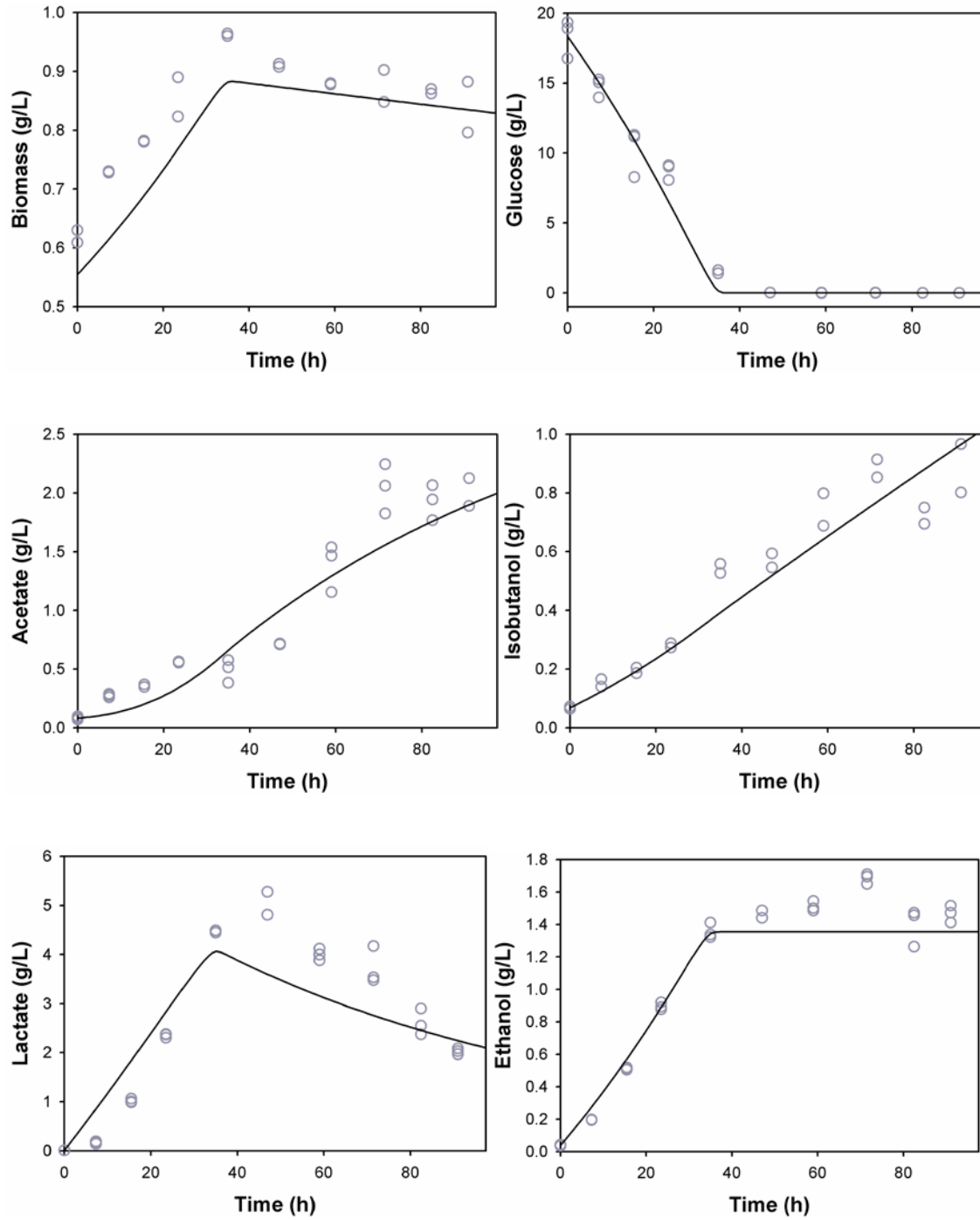


Figure 6.3: Growth kinetics after IPTG induction (F2).

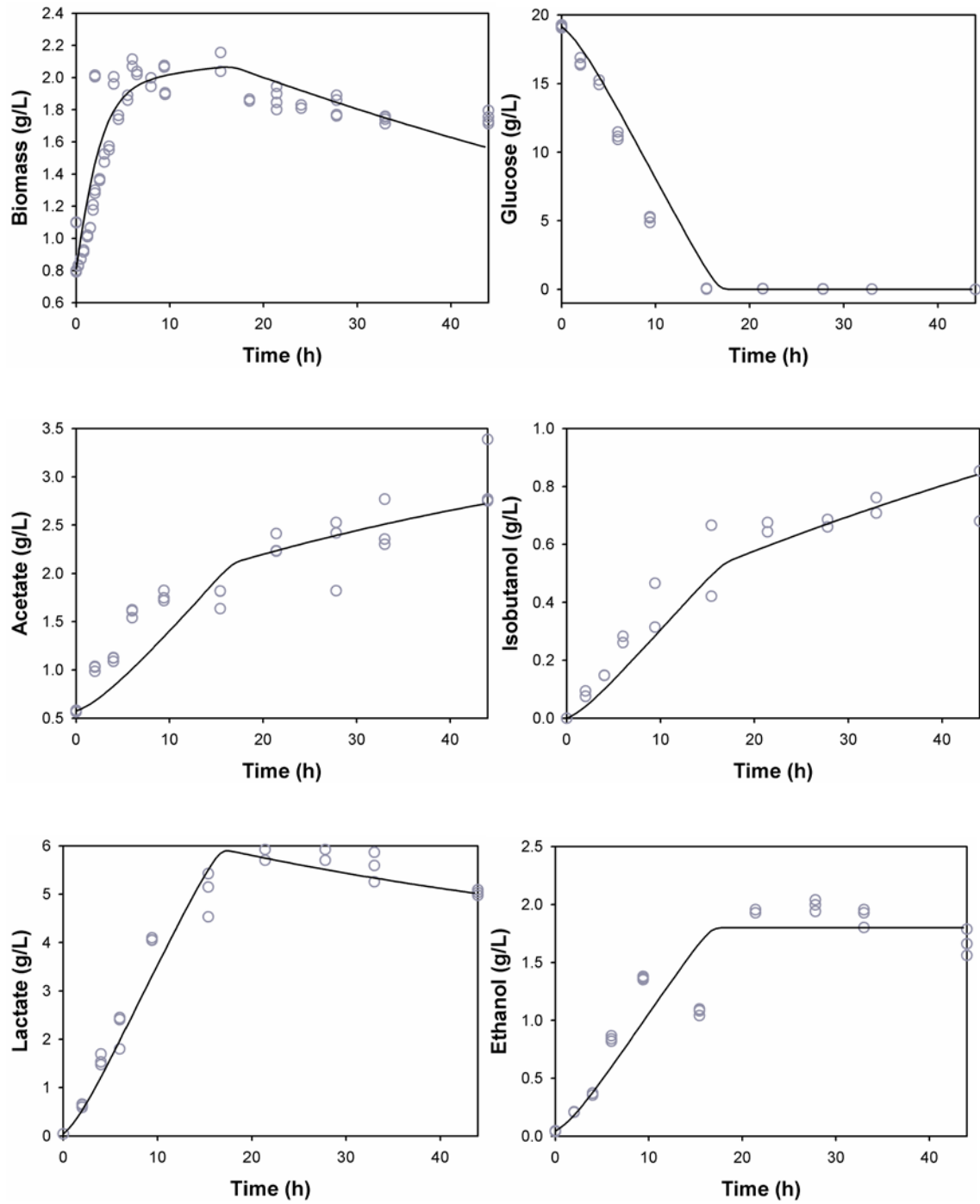


Figure 6.4: Growth kinetics after IPTG induction (F3, biomass growth data were from two identical batch experiments).

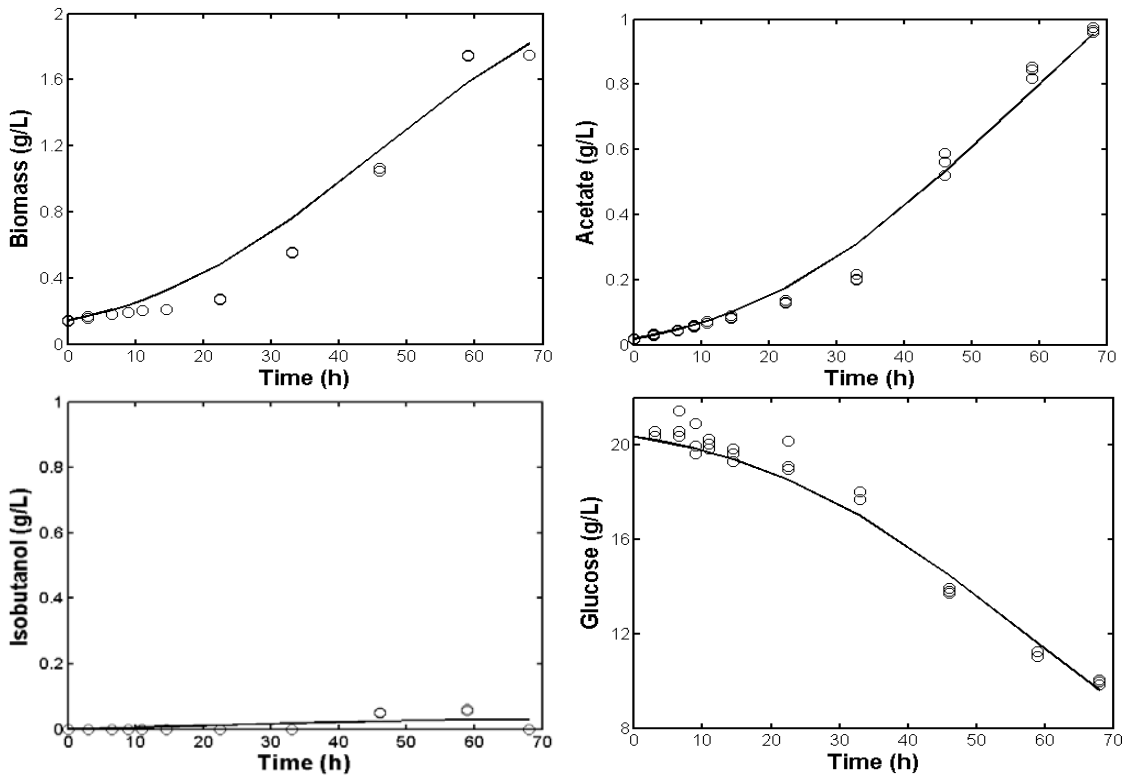


Figure 6.5: Growth kinetics after IPTG induction (F4).

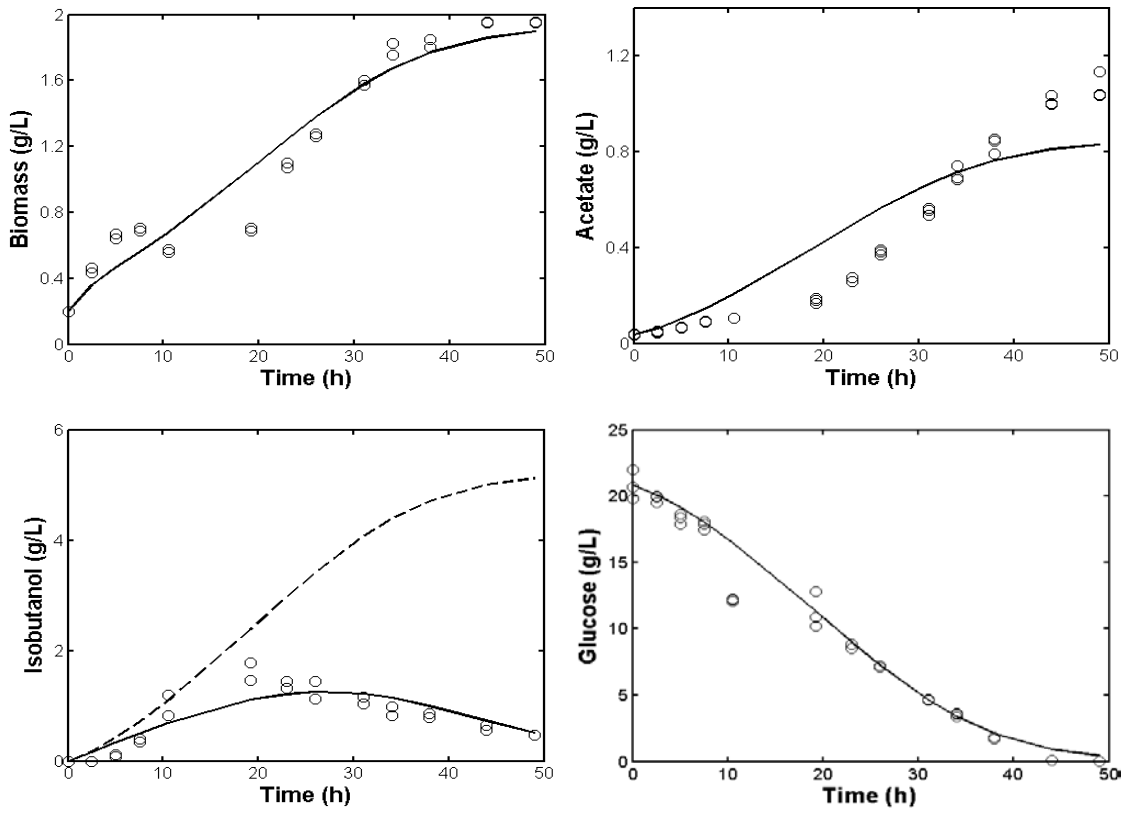


Figure 6.6: Growth kinetics after IPTG induction (F5). The dotted line was model prediction of IB concentrations without gas stripping.

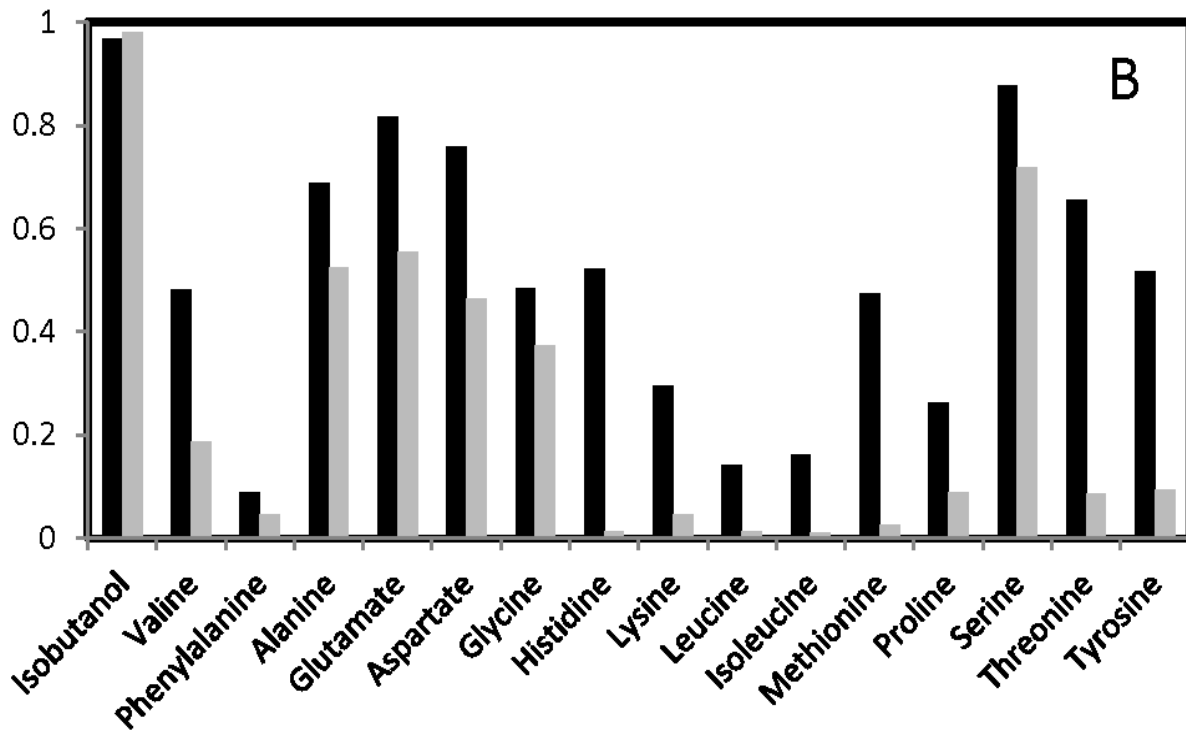
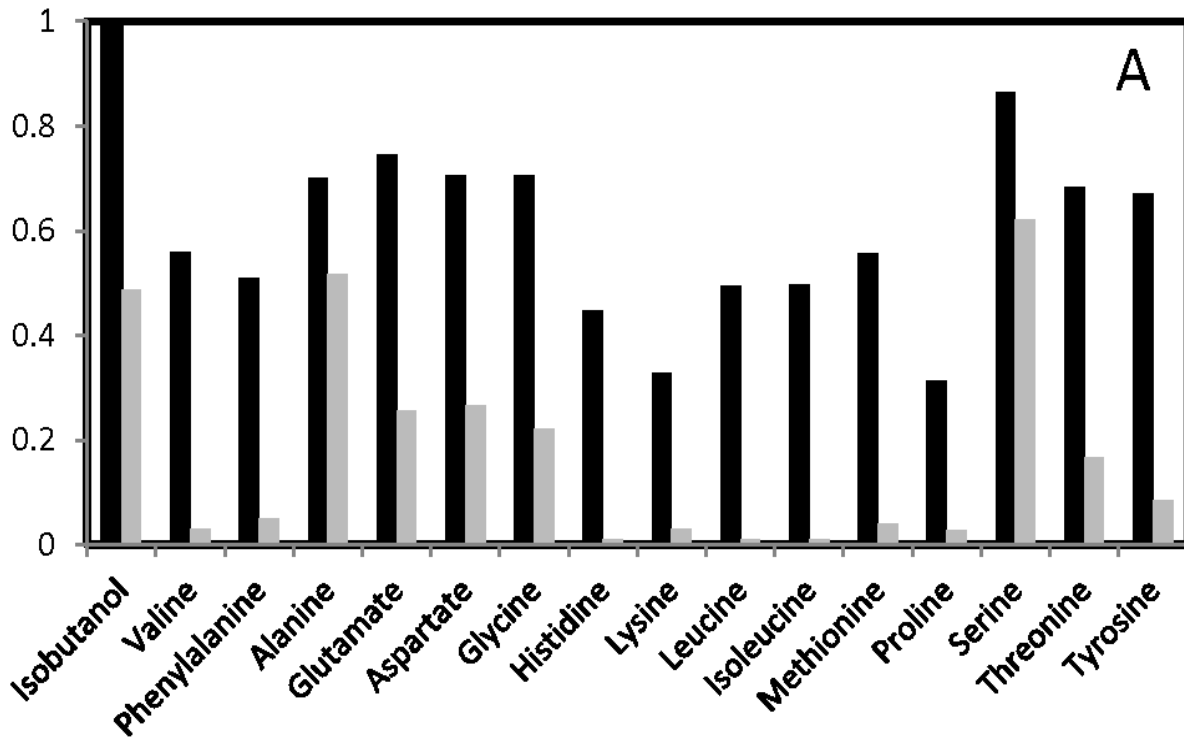


Figure 6.7: The fraction of ^{13}C carbon in metabolites from the low performance (A) and JCL260 (B) IB-producing strains. The biomass was grown on fully labeled ^{13}C -glucose, with 1 g/L (black bar) or 5 g/L (gray bar) nonlabeled yeast extract (n=2, GC-MS standard errors < 2%). The ^{13}C fractions (R) of metabolites were based on the following equation:

$$R = \frac{1}{n} \sum_{x=0}^n (x \cdot M_x)$$

where n was the total carbon number of the metabolite ($0 \leq x \leq n$). M_x was the corresponding ^{13}C isotopomer fraction for each isotopomer (M_0 was unlabeled fraction, M_1 was singly labeled fraction, M_2 was doubly labeled fraction, etc.)

Chapter 7: Conclusions and perspectives

7.1 Summary

With the advent of systems metabolic engineering, microbes have been engineered for synthesizing numerous chemicals and biofuels. However, there are still several roadblocks that remain to move microbial cell factories from laboratories to industry. In this thesis, we have performed several studies to overcome difficulties associated with development of efficient microbial platforms. In our first effort, we built a “Rule of Thumb” model to evaluate the various variables that influence microbial performance for the biosynthesis of diverse products under different growth conditions. Specifically, the yield of a microbial product remains difficult to calculate either by using the reaction stoichiometry or by using the large scale metabolic models. Filling this gap has been the focus for Chapter 2 and a statistical model was developed to get production yield of chemicals in *Saccharomyces Cerevisiae*. The developed statistical model allows the user to get a priori yield value based on the engineering to be performed. The model can also provide the degree of uncertainty associated with each parameter that can be used to improve yield of a product. As a second effort, the use of ^{13}C isotopomer analysis to elucidate the intrinsic product yields under complex nutrient conditions and multiple pathways for product synthesis has been discussed in Chapter 3. Moreover, in the same chapter we have also pointed out the value of ^{13}C -MFA in estimating the influence of microbial suboptimal energy metabolism on final product yield.

Besides modeling based studies, metabolic engineering tools were also applied to create three microbial platforms. Firstly, to contribute for the efforts in the field of biofuels and at the same time to reduce the dependence on food based biofuels, isobutanol synthesis from carbon

dioxide was demonstrated by engineering the cyanobacterium *Synechocystis* 6803. This research work also established the need for process integration along with metabolic engineering to improve microbial product titer. With the minimal engineering required for isobutanol synthesis, via co-metabolism of extra carbon substrates, and by using an *in situ* isobutanol removal system, we demonstrated improvements in isobutanol titer from the cyanobacterial platform. In the course of this work, isobutanol was found to be degraded photo chemically in the presence of hydroxyl radicals. This discovery necessitates research work for improvements in the reduction of radical accumulation during cell cultivation and thereby to reduce product degradation. With a view to offer industrial flexibility in handling carbon feedstock, mixotrophic fermentation was performed for isobutanol synthesis by providing the cultures with glucose. The growth of the strain did not increase as expected and the mechanism responsible for this counter action is likely due to metabolic imbalance during mixotrophic isobutanol production and it needs further investigation to elucidate the proper mechanism.

Secondly, the decreasing fossil fuel reserves will not only have its negative impact on the fuels but also negatively impact the petrochemicals that we use. PLA has been proposed as a substitute for polyethylene but presently its synthesis is food based. In Chapter 5, we have engineered the cyanobacterium *Synechocystis* 6803 to synthesize D-lactic acid. We have also demonstrated the positive effects of improving cofactor balance on the product titer along with improvements in the carbon flux. In this study, acetate was discovered to improve the photoautotrophic production of D-lactic acid by about two folds, possibly due to its inhibition of pyruvate decarboxylation reaction. By incorporating various metabolic engineering techniques and by feeding a cheap carbon feedstock, this work achieved the highest lactic acid titer ever to be reported using cyanobacteria as a host. In addition, this study also guided us in identifying the

target pathway (i.e., pyruvate decarboxylation) to improve microbial performance, which may be potentially regulated by utilizing synthetic circuits that use either a growth associated or a light activated promoter.

Thirdly, the kinetic models and the isotopomer studies developed in Chapter 6 allowed us to compare the metabolic performance of the two strains. This work enabled us to compare the role of nutrients in product synthesis between a low performance and a high performance strain, and to identify the rate limiting section of the biosynthesis pathway. Overall this thesis elaborates the combined application of isotopomer analysis, modeling and metabolic engineering research to improve microbial product yields.

7.2 Challenges with commercialization of industrial biotechnology

With the need to develop a sustainable technology for resolving environmental concerns, replacing fossil fuels and its derivatives, and creating new pharmaceutical chemicals for our better living, numerous metabolic engineering works were performed in the past decade. Despite the many successes that were attained in the laboratory, only a handful of them have reached commercialization. Listed below are some key reasons for the failure of metabolic engineering works to translate into microbial production at industrial scale ¹⁴⁹:

1. Compared to chemical synthesis, both the rate and the yield of microbial biosynthesis are very slow.
2. Substrate pretreatments are costly, reducing the profit margin for chemicals produced from microbial cell factories. For example, microbial hosts are still not efficient enough

in utilizing the low cost cellulose as their carbon feedstock, while the conversion of cellulose into sugars is still commercially challenging.

3. Product purification is also expensive if the fermentation titers are low.
4. Contamination associated with bioprocesses can lead to huge loss in product yield, but sterilization costs are very high.
5. Aerobic microbes need oxygen and the energy demand for intensive aeration makes bioprocessing very expensive. Moreover, enormous amounts of fresh water are required for fermentations.
6. Engineered strains are often unstable and therefore scaling up is very challenging.
7. Petroleum is still at an affordable price and therefore the profit margin for microbial productions is still very limited.

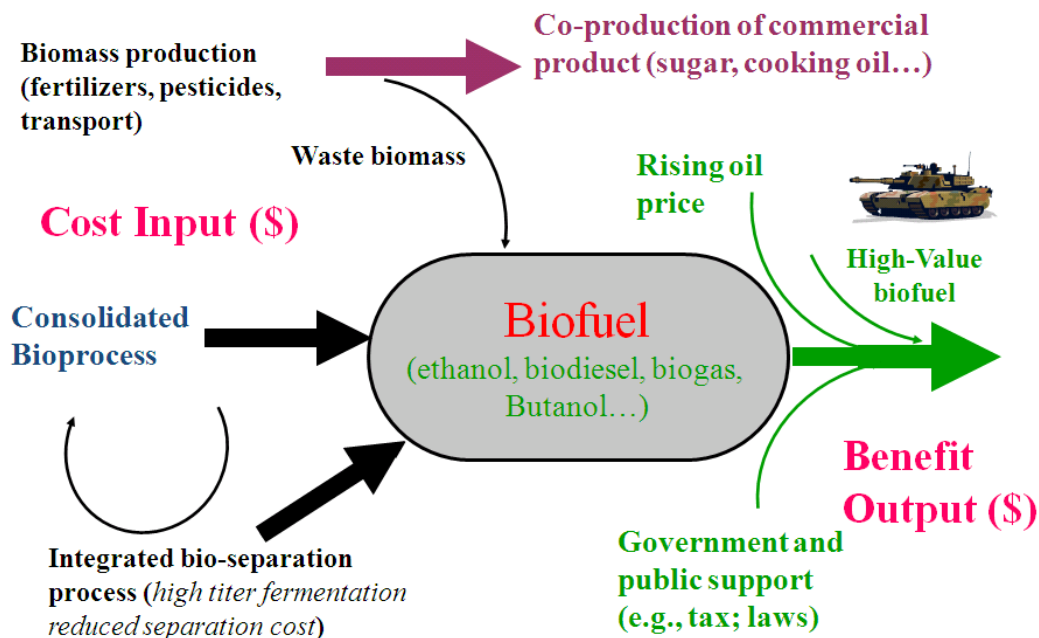


Figure 7.1: Schematic diagram of the various factors that play a major role in the economic feasibility of a biofuel production process.

In the past decade, several lessons were learnt from the biofuel industry and it cautions us that commercialization of biofuel production technology is a difficult task. Thereby, industrializing biofuel production would need the combined expertise of biologists, chemists, and engineers for its success. In order to reduce the cost of biofuel production, consolidated bioprocessing (remove the pretreatment step and integrated fermentation with bio-separation)²¹⁹ and co-production of value added products with biofuels are two key approaches in addition to metabolic engineering of microbial platforms (Figure 7.1). Moreover, global economy and government policy also influence the direction of biofuel industry. Therefore, we believe that engineering microbes for producing biofuels or other value-added chemicals is a promising direction that would require many years of hard work to realize its true potential.

7.3 Challenges with cyanobacterial bioprocessing

Metabolic engineering for the synthesis of value added products from cyanobacteria looks attractive as they can utilize carbon dioxide and sunlight. There are two main roadblocks for commercializing a technology that synthesizes products from engineered cyanobacteria. The first roadblock is related with cyanobacterial cultivation in large-scale. In large scale, cyanobacteria cultures are proposed to be grown either in open ponds or in closed photo-bioreactors. Although open pond are cheap for operation, they require year round sunlight as well as a warm climate, placing a strong limitation on the geographical location. Open ponds also have the other disadvantage of water loss by evaporation and a huge risk of microbial contamination. On the other hand, closed bioreactors looks like a good alternative, but they are still very expensive to operate ²²⁰. Algal photo-bioreactors also often suffer from higher

maintenance costs due to the formation of bio-films that block light penetrations. The second roadblock is associated with the low rates and titers of product synthesis in engineered cyanobacteria compared to their heterotrophic cousins. For example, an *E. coli* strain produced 5.2 g/L of fatty acids in 3 days²²¹, whereas the highest fatty acid titer in cyanobacteria was below 0.2 g/L after 2-weeks of cultivation¹⁷². Similarly, engineered *Synechocystis* 6803 achieved only 0.2 mg/L fatty alcohol, a 3000 fold lower titer as compared to the levels achieved by engineered *E. coli* (0.6 g/L)^{222, 30}. Cyanobacterial biosynthesis often takes weeks to synthesize a chemical in reasonable titers and this increases the operation along with maintenance costs. Long fermentations may also lead to product degradation and microbial contamination¹³⁶, further reducing the profit margin.

To overcome the cost due to the low productivity from cyanobacterial biofactories, we have proposed the integration of wastewater treatment with bio-production of D-lactate in our future research work. We believe that, the natural ability of cyanobacteria in utilizing N and P from wastewater along with its potential to synthesize value-added chemical synthesis from CO₂ using sunlight would result in a commercially viable process technology. Besides, strict life-cycle-analysis needs to be performed to reveal the energy, water and environmental impacts from these phototrophic microbial platforms²²³.

7.4 Recent developments in synthetic biology

Development of engineered microbes for artemisinic acid (precursor to the antimalarial drug artemisinin) production has been one of the major success stories since the inception of metabolic engineering⁷⁸. Towards developing a strain capable of synthesizing artemisinic acid at industrial levels²²⁴, it has been quoted that very little time and money were focused on

identifying the right metabolic pathways. The rest of the efforts were focused on applying the various synthetic biology tools to iteratively improve the performance of the strain ²²⁵. Synthetic biology was utilized to improve the production levels by carefully coordinating the expression of multiple genes to reduce intermediates accumulation, by balancing cofactors and by enriching precursors. Thereby, synthetic biology tools along with systems analysis and metabolic modeling can significantly speed up pathway optimization and strain development.

In recent studies, synthetic biology has extensively been applied for balancing metabolic pathways to increase the performance of the engineered microbes ²²⁶. The simplest strategy was to engineer cells by employing different promoters of varying strength and this has been successful on numerous occasions^{150, 227, 228}. Balancing the expression of all genes in a pathway can also be achieved by manipulating at mRNA level. This can be accomplished by varying the stability of specific mRNA segments that code for the enzyme ¹², by designing synthetic ribosome binding sites ²²⁹ and by utilizing transcription factor based approach to reprogram gene transcription at global level ⁸⁵. Metabolite channeling to improve product yield has also been performed to improve performance by utilizing synthetic protein scaffolds ^{230, 231}. Moreover, dynamic sensor-regulator systems have been developed to overcome toxicity of intermediates by switching the pathway at the correct time and thereby increasing the overall performance ¹¹⁰. Finally, high throughput genome engineering are being developed to speed up the creation of optimal hosts. For example, multiplexed automated genome engineering (MAGE), a strategy for large scale reprogramming of the genome based on natural selection principles may accelerate metabolic engineering by effectively tuning the expression of multiple genes ²³². In addition, trackable multiplex recombineering (TRMR) can create and evaluate thousands of genetic modifications simultaneously ²³³.

7.5 Future directions

Overall, there are still many challenges associated with developing industrial-strength strains for the synthesis of chemicals. However, several opportunities do exist to make microbial processes to be competitive with chemical synthesis. Along with developing more powerful synthetic biology tools, research must also focus on developing cost effective technologies to overcome challenges mentioned in Chapter 7.2. Thereby, microbial cell factories should be engineered to metabolize multiple substrates, synthesize multiple products, reduce byproducts, and to minimize oxygen demand as well as to engineer them with control systems to uptake key precursors and synthesize products as needed. All these engineering must be performed in an integrated manner with a systems level understanding of the cell at all levels (Figure 7.2). Systems level analysis will enable researchers to identify bottleneck pathways and genes that can be targeted for improved performance. Model development to simulate the output of synthetic biology tools would enable us to predict and understand the dynamics of engineered pathways. This development would lead to a tremendous reduction in experimental hours by screening for optimal pathway designs on a computer, before engineering it into a host cell. Also, to improve the economical margin of cyanobacterial based product synthesis, the engineered algal process can be integrated with a wastewater treatment facility. On the other hand, our incomplete knowledge about the biology of the cell often requires guesses to perform metabolic engineering. As our understanding about the cell grows, guesswork based experiments would be avoided leading to more successful outcomes. Simultaneously, we have to work with process engineers to bridge gaps between laboratories studies and industrial applications. With all these

developments, it is possible that in the near future, we may realize the dream of using microbial cell factories for the production of diverse value-added chemicals at industrial scale from cheaper feedstock.

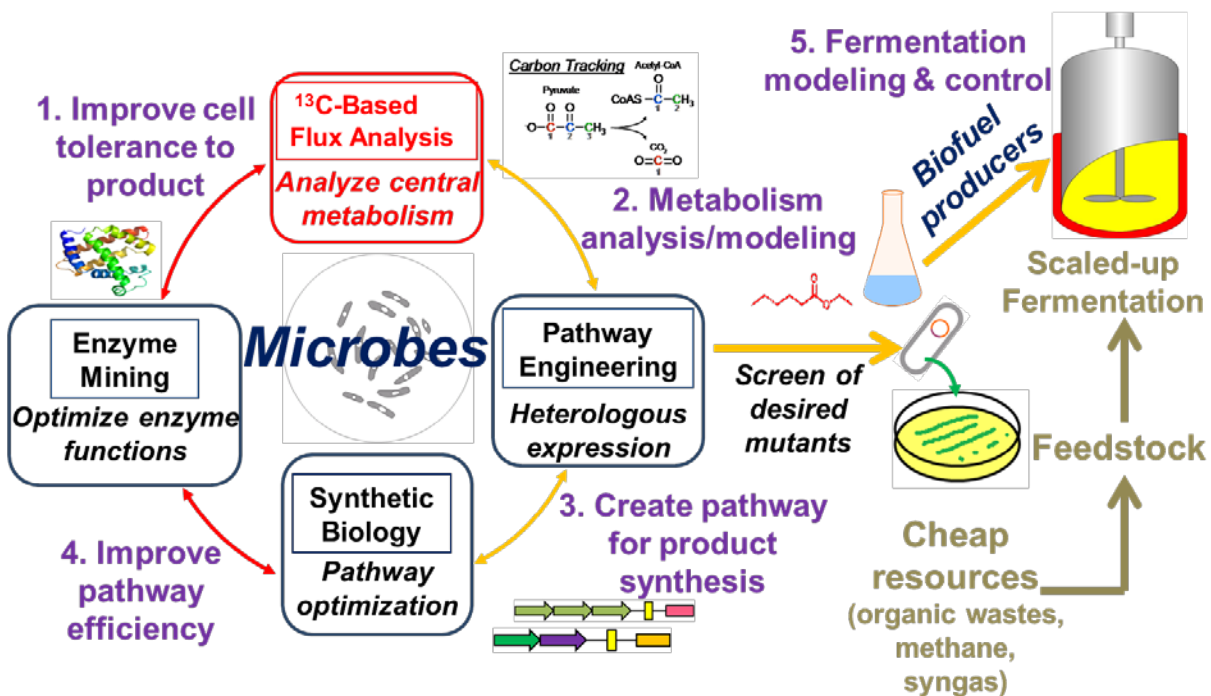


Figure 7.2: Schematic diagram of an integrated iterative approach required for the development of high-performance microbial strains towards industrial commercialization.

References

1. Zhou, J. & Li, Y. Engineering cyanobacteria for fuels and chemicals production. *Protein & Cell* **1**, 207-210 (2010).
2. Administration, U.S.E.I., Vol. 2011 (U.S. Energy Information Administration, Washington; 2010).
3. Agency, U.S.E.P., Vol. 2011 (U.S. Environmental Protection Agency, Washington; 2010).
4. Jarboe, L.R. et al. Metabolic engineering for production of biorenewable fuels and chemicals: contributions of synthetic biology. *Journal of Biomedicine and Biotechnology* (2010).
5. Bailey, J. Toward a science of metabolic engineering. *Science (New York, NY)* **252**, 1668 - 1675 (1991).
6. Lee, S.Y., Kim, H.U., Park, J.H., Park, J.M. & Kim, T.Y. Metabolic engineering of microorganisms: general strategies and drug production. *Drug Discovery Today* **14**, 78-88 (2009).
7. Chemler, J., Yan, Y. & Koffas, M. Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. *Microbial Cell Factories* **5**, 20 (2006).
8. Alper, H. & Stephanopoulos, G. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology* **7**, 715-723 (2009).
9. Clomburg, J.M. & Gonzalez, R. Biofuel production in *Escherichia coli*: the role of metabolic engineering and synthetic biology. *Applied Microbiology and Biotechnology* **86**, 419-434 (2010).
10. Higashide, W., Li, Y., Yang, Y. & Liao, J.C. Metabolic Engineering of *Clostridium cellulolyticum* for Production of Isobutanol from Cellulose. *Applied and Environmental Microbiology* **77**, 2727-2733 (2011).
11. Steen, E.J. et al. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-562 (2010).
12. Pfleger, B.F., Pitera, D.J., Smolke, C.D. & Keasling, J.D. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature biotechnology* **24**, 1027-1032 (2006).
13. Ducat, D.C., Way, J.C. & Silver, P.A. Engineering cyanobacteria to generate high-value products. *Trends in Biotechnology* **29**, 95-103 (2011).
14. Hu, Q. et al. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal* **54**, 621-639 (2008).
15. Kaneko, T. et al. Structural analysis of four large plasmids harboring in a unicellular cyanobacterium, *Synechocystis* sp. PCC 6803. *DNA Res* **10**, 221-228 (2003).
16. Kaneko, T. et al. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res* **3**, 109-136 (1996).
17. Vermaas, W. Molecular genetics of the cyanobacterium *Synechocystis* sp. PCC 6803: Principles and possible biotechnology applications. *Journal of Applied Phycology* **8**, 263-273 (1996).

18. Los, D.A. et al. Stress Sensors and Signal Transducers in Cyanobacteria. *Sensors* **10**, 2386-2415 (2010).
19. Zang, X.N., Liu, B., Liu, S.M., Arunakumara, K.K.I.U. & Zhang, X.C. Optimum conditions for transformation of *Synechocystis* sp PCC 6803. *Journal of Microbiology* **45**, 241-245 (2007).
20. Devroe, E.J.M.M.A. et al., Vol. US 7,785,861 B2 (Joule Unlimited, US; 2010).
21. Dexter, J. & Fu, P. Metabolic engineering of cyanobacteria for ethanol production. *Energy & Environmental Science* **2**, 857-864 (2009).
22. Biofuels, A., Vol. 2011 (Algenol Biofuels, 2009).
23. Steen, E.J. et al. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microbial Cell Factories* **7**, 36 (2008).
24. Shen, C.R. et al. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl Environ Microbiol* **77**, 2905-2915 (2011).
25. Atsumi, S., Hanai, T. & Liao, J.C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-89 (2008).
26. Sentheshanmuganathan, S. & Elsdén, S.R. The mechanism of the formation of tyrosol by *Saccharomyces cerevisiae*. *The Biochemical journal* **69**, 210-218 (1958).
27. Hazelwood, L.A., Daran, J.M., van Maris, A.J., Pronk, J.T. & Dickinson, J.R. The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. *Appl Environ Microbiol* **74**, 2259-2266 (2008).
28. Atsumi, S., Higashide, W. & Liao, J.C. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nature biotechnology* **27**, 1177-1180 (2009).
29. Lan, E.I. & Liao, J.C. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metabolic engineering* **13**, 353-363 (2011).
30. Tan, X. et al. Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metabolic engineering* **13**, 169-176 (2011).
31. Schirmer, A., Rude, M.A., Li, X., Popova, E. & del Cardayre, S.B. Microbial Biosynthesis of Alkanes. *Science* **329**, 559-562 (2010).
32. Ducat, D.C., Sachdeva, G. & Silver, P.A. Rewiring hydrogenase-dependent redox circuits in cyanobacteria. *Proceedings of the National Academy of Sciences* **108**, 3941-3946 (2011).
33. Bailey, J.E. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress* **14**, 8-20 (1998).
34. Stephanopoulos, N., Aristidou, A. & Nielsen, J. *Metabolic engineering principles and methodologies*. (Academic Press, San Diego; 1998).
35. Krivoruchko, A., Siewers, V. & Nielsen, J., Vol. 6 262-276 (WILEY-VCH Verlag, 2011).
36. Clark, D.S. & Blanch, H.W. *Biochemical engineering*, Edn. 1st. (CRC Press, Boca Raton; 1997).
37. Pelczar Jr., M.J., Chan, E.C.S. & Krieg, N.R. *Microbiology*, Edn. 5th. (McGraw-Hill, New York; 1998).
38. Hahn-Hagerdal, B. et al. Role of cultivation media in the development of yeast strains for large scale industrial use. *Microbial Cell Factories* **4**, 31 (2005).
39. James E. Bailey & Ollis, D.F. (eds.) *Biochemical engineering fundamentals*, Edn. 2. (McGraw-Hill, New York; 1986).

40. Weisberg, S. Applied Linear Regression, Edn. 2nd. (John Wiley & Sons, New York; 1985).
41. Tang, Y.J., Qi, L. & Krieger-Brockett, B. Evaluating factors that influence microbial phenanthrene biodegradation rates by regression with categorical variables. *Chemosphere* **59**, 729-741 (2005).
42. Team, R.D.C. (R Foundation for Statistical Computing, Vienna; 2010).
43. du Prel, J.B., Hommel, G., Rohrig, B. & Blettner, M. Confidence interval or p-value? part 4 of a series on evaluation of scientific publications. *Deutsches Arzteblatt International* **106**, 335-339 (2009).
44. Jiang, H.X., Wood, K.V. & Morgan, J.A. Metabolic engineering of the phenylpropanoid pathway in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **71**, 2962-2969 (2005).
45. Maltsev, N., Glass, E.M., Ovchinnikova, G. & Gu, Z. Molecular Mechanisms Involved in Robustness of Yeast Central Metabolism against Null Mutations. *Journal of Biochemistry* **137**, 177-187 (2005).
46. Leonard, E. et al. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13654-13659 (2010).
47. Woolston, B.M., Edgar, S. & Stephanopoulos, G. Metabolic engineering: past and future. *Annual review of chemical and biomolecular engineering* **4**, 259-288 (2013).
48. Merico, A., Sulo, P., Piskur, J. & Compagno, C. Fermentative lifestyle in yeasts belonging to the *Saccharomyces* complex. *FEBS Journal* **274**, 976-989 (2007).
49. Piskur, J., Rozpedowska, E., Polakova, S., Merico, A. & Compagno, C. How did *Saccharomyces* evolve to become a good brewer? *Trends in Genetics* **22**, 183-186 (2006).
50. Bro, C., Regenberg, B., Forster, J. & Nielsen, J. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic engineering* **8**, 102-111 (2006).
51. Wen, F., Sun, J. & Zhao, H.M. Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol. *Applied and Environmental Microbiology* **76**, 1251-1260 (2009).
52. Madsen, K.M. et al. Linking Genotype and Phenotype of *Saccharomyces cerevisiae* Strains Reveals Metabolic Engineering Targets and Leads to Triterpene Hyper-Producers. *PLoS ONE* **6**, e14763 (2011).
53. Fossati, T., Solinas, N., Porro, D. & Branduardi, P. l-ascorbic acid producing yeasts learn from plants how to recycle it. *Metabolic engineering* **13**, 177-185 (2011).
54. Rico, J., Pardo, E. & Orejas, M. Enhanced production of a plant monoterpene by overexpression of the 3-hydroxy-3-methylglutaryl coenzyme A reductase catalytic domain in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **76**, 6449-6454 (2010).
55. Brochado, A. et al. Improved vanillin production in baker's yeast through in silico design. *Microbial Cell Factories* **9**, 84 (2010).
56. Raab, A.M., Gebhardt, G., Bolotina, N., Weuster-Botz, D. & Lang, C. Metabolic engineering of *Saccharomyces cerevisiae* for the biotechnological production of succinic acid. *Metabolic engineering* **12**, 518-525 (2010).

57. Dmytruk, K.V., Yatsyshyn, V.Y., Sybirna, N.O., Fedorovych, D.V. & Sibirny, A.A. Metabolic engineering and classic selection of the yeast *Candida famata* (*Candida flareri*) for construction of strains with enhanced riboflavin production. *Metabolic engineering* **13**, 82-88 (2011).
58. Hu, X.Q. et al. Effects of different glycerol feeding strategies on S-adenosyl-L-methionine biosynthesis by P-GAP-driven *Pichia pastoris* overexpressing methionine adenosyltransferase. *Journal of Biotechnology* **137**, 44-49 (2008).
59. Ilmen, M., Koivuranta, K., Ruohonen, L., Suominen, P. & Penttila, M. Efficient production of L-lactic acid from xylose by *Pichia stipitis*. *Applied and Environmental Microbiology* **73**, 117-123 (2007).
60. Popp, A. et al. Fermentative production of L-glycerol 3-phosphate utilizing a *Saccharomyces cerevisiae* strain with an engineered glycerol biosynthetic pathway. *Biotechnology and Bioengineering* **100**, 497-505 (2008).
61. Zhuo, R.Y. et al. Identification and Characterization of Novel MicroRNAs from *Populus cathayana* Rehd. *Plant Molecular Biology Reporter* **29**, 242-251 (2011).
62. Zelle, R.M. et al. Malic acid production by *Saccharomyces cerevisiae*: Engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. *Applied and Environmental Microbiology* **74**, 2766-2777 (2008).
63. Colletti, P. et al. Evaluating factors that influence microbial synthesis yields by linear regression with numerical and ordinal variables. *Biotechnology and Bioengineering* **108**, 893-901 (2011).
64. Farmer, W.R. & Liao, J.C. Precursor balancing for metabolic engineering of lycopene production in *Escherichia coli*. *Biotechnology Progress* **17**, 57-61 (2001).
65. Blank, L.M., Kuepfer, L. & Sauer, U. Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome biology* **6**, R49 (2005).
66. Kauffman, K.J., Prakash, P. & Edwards, J.S. Advances in flux balance analysis. *Current Opinion in Biotechnology* **14**, 491-496 (2003).
67. VanBriesen, J.M. Evaluation of methods to predict bacterial yield using thermodynamics. *Biodegradation* **13**, 171-190 (2002).
68. Marx, H., Mattanovich, D. & Sauer, M. Overexpression of the riboflavin biosynthetic pathway in *Pichia pastoris*. *Microbial Cell Factories* **7** (2008).
69. Bao, W.G. et al. Oxygen-dependent transcriptional regulator Hap1p limits glucose uptake by repressing the expression of the major glucose transporter gene RAG1 in *Kluyveromyces lactis*. *Eukaryotic Cell* **7**, 1895-1905 (2008).
70. Bhataya, A., Schmidt-Dannert, C. & Lee, P.C. Metabolic engineering of *Pichia pastoris* X-33 for lycopene production. *Process Biochemistry* **44**, 1095-1102 (2009).
71. Tokuhiro, K., Muramatsu M et al. Overproduction of geranylgeraniol by metabolically engineered *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **75**, 5536-5543 (2009).
72. Lee, W. & DaSilva, N.A. Application of sequential integration for metabolic engineering of 1,2-propanediol production in yeast. *Metabolic engineering* **8**, 58-65 (2006).
73. Jeon, E. et al. Development of a *Saccharomyces cerevisiae* strain for the production of 1,2-propanediol by gene manipulation. *Enzyme and Microbial Technology* **45**, 42-47 (2009).

74. Takahashi, S. et al. Metabolic engineering of sesquiterpene metabolism in yeast. *Biotechnology and Bioengineering* **97**, 170-181 (2007).
75. Cordier, H., Mendes, F., Vasconcelos, I. & Francois, J.M. A metabolic and genomic study of engineered *Saccharomyces cerevisiae* strains for high glycerol production. *Metabolic engineering* **9**, 364-378 (2007).
76. Van Vleet, J.H., Jeffries, T.W. & Olsson, L. Deleting the para-nitrophenyl phosphatase (pNPPase), PHO13, in recombinant *Saccharomyces cerevisiae* improves growth and ethanol production on d-xylose. *Metabolic engineering* **10**, 360-369 (2008).
77. Sonderegger, M., Schumperli, M. & Sauer, U. Metabolic engineering of a phosphoketolase pathway for pentose catabolism in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **70**, 2892-2897 (2004).
78. Ro, D.K., Paradise EM, Ouellet M, Fisher K J, Newman K L, Ndungu J M et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943 (2006).
79. Shiba, Y., Paradise, E.M., Kirby, J., Ro, D.K. & Keasling, J.D. Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metabolic engineering* **9**, 160-168 (2007).
80. Steinle, A., Bergander, K. & Steinbuchel, A. Metabolic engineering of *Saccharomyces cerevisiae* for production of novel cyanophycins with an extended range of constituent amino acids. *Applied and Environmental Microbiology* **75**, 3437-3446 (2009).
81. Nguyen, H.T.T. & Nevoigt, E. Engineering of *Saccharomyces cerevisiae* for the production of dihydroxyacetone (DHA) from sugars: A proof of concept. *Metabolic engineering* **11**, 335-346 (2009).
82. Ishida, N. et al. Metabolic engineering of *Saccharomyces cerevisiae* for efficient production of pure L-(+)- lactic acid. *Applied Biochemistry and Biotechnology* **131**, 795-807 (2006).
83. Szkopinska, A., Swiezewska, E. & Karst, F. The regulation of activity of main mevalonic acid pathway enzymes: farnesyl diphosphate synthase, 3-hydroxy-3-methylglutaryl-CoA reductase, and squalene synthase in yeast *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* **267**, 473-477 (2000).
84. Toivari, M.H., Aristidou, A., Ruohonen, L. & Penttila, M. Conversion of xylose to ethanol by recombinant *Saccharomyces cerevisiae*: Importance of xylulokinase (XKS1) and oxygen availability. *Metabolic engineering* **3**, 236-249 (2001).
85. Alper, H., Moxley, J., Nevoigt, E., Fink, G.R. & Stephanopoulos, G. Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production. *Science* **314**, 1565-1568 (2006).
86. Hong, M.-E. et al. Identification of gene targets eliciting improved alcohol tolerance in *Saccharomyces cerevisiae* through inverse metabolic engineering. *Journal of Biotechnology* **149**, 52-59 (2010).
87. Pirkov, I., Albers, E., Norbeck, J. & Larsson, C. Ethylene production by metabolic engineering of the yeast *Saccharomyces cerevisiae*. *Metabolic engineering* **10**, 276-280 (2008).
88. Yan, Y., Kohli, A. & Koffas, M.A. Biosynthesis of natural flavanones in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **71**, 5610-5613 (2005).

89. Kennedy, C.J., Boyle, P.M., Waks, Z. & Silver, P.A. Systems-level engineering of nonfermentative metabolism in yeast. *Genetics* **183**, 385-397 (2009).
90. Oswald, M., Fischer, M., Dirninger, N. & Karst, F. Monoterpenoid biosynthesis in *Saccharomyces cerevisiae*. *FEMS Yeast Research* **7**, 413-421 (2007).
91. Remize, F., Barnavon, L. & Dequin, S. Glycerol export and glycerol-3-phosphate dehydrogenase, but not glycerol phosphatase, are rate limiting for glycerol production in *Saccharomyces cerevisiae*. *Metabolic engineering* **3**, 301-312 (2001).
92. Overkamp, K.M. et al. Metabolic engineering of glycerol production in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **68**, 2814-2821 (2002).
93. Geertman, J.-M.A., van Maris, A.J.A., van Dijken, J.P. & Pronk, J.T. Physiological and genetic engineering of cytosolic redox metabolism in *Saccharomyces cerevisiae* for improved glycerol production. *Metabolic engineering* **8**, 532-542 (2006).
94. Szczebara, F.M., Chandelier C, Villeret C, Masurel A, Bourot S, Duport C, Blanchard S, Groisillier A, Testet E et al. Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nature biotechnology* **21**, 143-149 (2003).
95. Skory, C.D. Lactic acid production by *Saccharomyces cerevisiae* expressing a *Rhizopus oryzae* lactate dehydrogenase gene. *Journal of Industrial Microbiology and Biotechnology* **30**, 22-27 (2003).
96. Colombie, S. & Sablayrolles, J.M. Nicotinic acid controls lactate production by K1-LDH: a *Saccharomyces cerevisiae* strain expressing a bacterial LDH gene. *Journal of Industrial Microbiology and Biotechnology* **31**, 209-215 (2004).
97. Sauer, M., Branduardi, P., Valli, M. & Porro, D. Production of L-ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. *Applied and Environmental Microbiology* **70**, 6086-6091 (2004).
98. Ishida, N. et al. Efficient production of L-lactic acid by metabolically engineered *Saccharomyces cerevisiae* with a genome-integrated L-lactate dehydrogenase gene. *Applied and Environmental Microbiology* **71**, 1964-1970 (2005).
99. Trantas, E., Panopoulos, N. & Ververidis, F. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metabolic engineering* **11**, 355-366 (2009).
100. Vannelli, T., Qi, W.W., Sweigard, J., Gatenby, A.A. & Sariaslani, F.S. Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metabolic engineering* **9**, 142-151 (2007).
101. Carlson, R. & Srienc, F. Effects of recombinant precursor pathway variations on poly[(R)-3-hydroxybutyrate] synthesis in *Saccharomyces cerevisiae*. *Journal of Biotechnology* **124**, 561-573 (2006).
102. van Maris, A.J.A. et al. Directed evolution of Pyruvate Decarboxylase-negative *Saccharomyces cerevisiae*, yielding a C2-independent, glucose-tolerant, and pyruvate-hyperproducing Yeast. *Applied and Environmental Microbiology* **70**, 159-166 (2004).
103. Hawkins, K.M. & Smolke, C.D. Production of benzyloquinoline alkaloids in *Saccharomyces cerevisiae*. *Nature chemical biology* **4**, 564-573 (2008).
104. Toivari, M.H., Ruohonen, L., Miasnikov, A.N., Richard, P. & Penttila, M. Metabolic engineering of *Saccharomyces cerevisiae* for conversion of D-glucose to xylitol and other

- five-carbon sugars and sugar alcohols. *Applied and Environmental Microbiology* **73**, 5471-5476 (2007).
105. DeJong, J.M. et al. Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering* **93**, 212-224 (2006).
 106. Hansen, E.H. et al. De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Applied and Environmental Microbiology* **75**, 2765-2774 (2009).
 107. Verwaal, R. et al. High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous*. *Applied and Environmental Microbiology* **73**, 4342-4350 (2007).
 108. Peralta-Yahya, P.P., Zhang, F., del Cardayre, S.B. & Keasling, J.D. Microbial engineering for the production of advanced biofuels. *Nature* **488**, 320-328 (2012).
 109. Dellomonaco, C., Clomburg, J.M., Miller, E.N. & Gonzalez, R. Engineered reversal of the β -oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**, 355-359 (2011).
 110. Zhang, F., Carothers, J.M. & Keasling, J.D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat Biotech* **30**, 354-359 (2012).
 111. Choi, Y.J. & Lee, S.Y. Microbial production of short-chain alkanes. *Nature advance online publication* (2013).
 112. Xu, P. et al. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nat Commun* **4**, 1409 (2013).
 113. Keasling, J.D., Mendoza, A. & Baran, P.S. Synthesis: A constructive debate. *Nature* **492**, 188-189 (2012).
 114. Bastian, S. et al. Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. *Metabolic engineering* **13**, 345-352 (2011).
 115. Huffer, S., Roche, C.M., Blanch, H.W. & Clark, D.S. *Escherichia coli* for biofuel production: bridging the gap from promise to practice. *Trends Biotechnol* **30**, 538-545 (2012).
 116. Stephanopoulos, G. Challenges in Engineering Microbes for Biofuels Production. *Science* **315**, 801-804 (2007).
 117. Bogorad, I.W., Lin, T.-S. & Liao, J.C. Synthetic non-oxidative glycolysis enables complete carbon conservation. *Nature* doi:10.1038/nature12575 (2013).
 118. Bond-Watts, B.B., Bellerose, R.J. & Chang, M.C. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nature chemical biology* **7**, 222-227 (2011).
 119. Tang, J.K.-H., You, L., Blankenship, R.E. & Tang, Y.J. Recent advances in mapping environmental microbial metabolisms through ¹³C isotopic fingerprints. *Journal of The Royal Society Interface* **9**, 2767-2780 (2012).
 120. Feng, X. et al. Characterization of the Central Metabolic Pathways in *Thermoanaerobacter* sp. Strain X514 via Isotopomer-Assisted Metabolite Analysis. *Applied and Environmental Microbiology* **75**, 5001-5008 (2009).
 121. Misra, A. et al. Metabolic analyses elucidate nontrivial gene targets for amplifying dihydroartemisinic acid production in yeast. *Frontiers in Microbiology* **4** (2013).

122. Suthers, P.F. et al. Metabolic flux elucidation for large-scale models using ¹³C labeled isotopes. *Metabolic engineering* **9**, 387-405 (2007).
123. Stephanopoulos, G. Metabolic Fluxes and Metabolic Engineering. *Metabolic engineering* **1**, 1-11 (1999).
124. Zamboni, N., Fendt, S.M., Rühl, M. & Sauer, U. ¹³C-based metabolic flux analysis. *Nature Protocols* **4**, 878-892 (2009).
125. Zamboni, N. & Sauer, U. Novel biological insights through metabolomics and ¹³C-flux analysis. *Current Opinion in Microbiology* **12**, 553-558 (2009).
126. Wiechert, W. ¹³C Metabolic Flux Analysis. *Metabolic engineering* **3**, 195-206 (2001).
127. Antoniewicz, M.R. et al. Metabolic flux analysis in a nonstationary system: fed-batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. *Metabolic engineering* **9**, 277-292 (2007).
128. Varman, A.M., Xiao, Y., Leonard, E. & Tang, Y. Statistics-based model for prediction of chemical biosynthesis yield from *Saccharomyces cerevisiae*. *Microbial Cell Factories* **10**, 45 (2011).
129. Xiao, Y. et al. Kinetic Modeling and Isotopic Investigation of Isobutanol Fermentation by Two Engineered *Escherichia coli* Strains. *Industrial & Engineering Chemistry Research* **51**, 15855-15863 (2012).
130. Baez, A., Cho, K.-M. & Liao, J.C. High-flux isobutanol production using engineered *Escherichia coli*: a bioreactor study with in situ product removal. *Applied Microbiology and Biotechnology* **90**, 1681-1690 (2011).
131. Xiao, Y. et al. Engineering *Escherichia coli* to convert acetic acid to free fatty acids. *Biochemical Engineering Journal* **76**, 60-69 (2013).
132. Yang, C., Hua, Q. & Shimizu, K. Metabolic Flux Analysis in *Synechocystis* Using Isotope Distribution from ¹³C-Labeled Glucose. *Metabolic engineering* **4**, 202-216 (2002).
133. Varman, A.M., Yu, Y., You, L. & Tang, Y.J. Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microbial Cell Factories* **12**, 117 (2013).
134. Goodson, C., Roth, R., Wang, Z.T. & Goodenough, U. Structural Correlates of Cytoplasmic and Chloroplast Lipid Body Synthesis in *Chlamydomonas reinhardtii* and Stimulation of Lipid Body Production with Acetate Boost. *Eukaryotic Cell* **10**, 1592-1606 (2011).
135. Yu, Y. et al. Development of *Synechocystis* sp. PCC 6803 as a Phototrophic Cell Factory. *Marine Drugs* **11**, 2894-2916 (2013).
136. Varman, A.M., Xiao, Y., Pakrasi, H.B. & Tang, Y.J. Metabolic Engineering of *Synechocystis* sp. Strain PCC 6803 for Isobutanol Production. *Applied and Environmental Microbiology* **79**, 908-914 (2013).
137. Rungtaphan, W. & Keasling, J.D. Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid-derived biofuels and chemicals. *Metabolic engineering* <http://dx.doi.org/10.1016/j.ymben.2013.07.003> (2013).
138. Varma, A. & Palsson, B.O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and Environmental Microbiology* **60**, 3724-3731 (1994).

139. Mitsumori, F., Rees, D., Brindle, K.M., Radda, G.K. & Campbell, I.D. 31P-NMR saturation transfer studies of aerobic *Escherichia coli* cells. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **969**, 185-193 (1988).
140. Sauer, U. & Bailey, J.E. Estimation of P-to-O ratio in *Bacillus subtilis* and its influence on maximum riboflavin yield. *Biotechnology and Bioengineering* **64**, 750-754 (1999).
141. Varela, C.A., Baez, M.E. & Agosin, E. Osmotic stress response: quantification of cell maintenance and metabolic fluxes in a lysine-overproducing strain of *Corynebacterium glutamicum*. *Appl Environ Microbiol* **70**, 4222-4229 (2004).
142. Lennen, R.M. & Pfleger, B.F. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends Biotechnol* **30**, 659-667 (2012).
143. He, L. et al. Central metabolic responses to the overproduction of fatty acids in *Escherichia coli* based on ¹³C-metabolic flux analysis. *Biotechnology and Bioengineering* doi: **10.1002/bit.25124** (2013).
144. Lennen, R.M. et al. Membrane stresses induced by overproduction of free fatty acids in *Escherichia coli*. *Appl Environ Microbiol* **77**, 8114-8128 (2011).
145. Feng, X. & Zhao, H. Investigating xylose metabolism in recombinant *Saccharomyces cerevisiae* via ¹³C metabolic flux analysis. *Microbial Cell Factories* **12**, 114 (2013).
146. Liou, J.S.-C., Balkwill, D.L., Drake, G.R. & Tanner, R.S. *Clostridium carboxidivorans* sp. nov., a solvent-producing clostridium isolated from an agricultural settling lagoon, and reclassification of the acetogen *Clostridium scatologenes* strain SL1 as *Clostridium drakei* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **55**, 2085-2091 (2005).
147. Mattozzi, M.d., Ziesack, M., Voges, M.J., Silver, P.A. & Way, J.C. Expression of the sub-pathways of the *Chloroflexus aurantiacus* 3-hydroxypropionate carbon fixation bicycle in *E. coli*: Toward horizontal transfer of autotrophic growth. *Metabolic engineering* **16**, 130-139 (2013).
148. Atsumi, S. & Liao, J. Directed evolution of *Methanococcus jannaschii* citramalate synthase for biosynthesis of 1-propanol and 1-butanol by *Escherichia coli*. *Appl Environ Microbiol* **74**, 7802 - 7808 (2008).
149. Chen, G.-Q. New challenges and opportunities for industrial biotechnology. *Microbial Cell Factories* **11**, 111 (2012).
150. Ajikumar, P.K. et al. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **330**, 70-74 (2010).
151. Bennett, B.D., Yuan, J., Kimball, E.H. & Rabinowitz, J.D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat. Protocols* **3**, 1299-1311 (2008).
152. Rittmann, B.E. Opportunities for renewable bioenergy using microorganisms. *Biotechnology and Bioengineering* **100**, 203-212 (2008).
153. van Groenigen, K.J., Osenberg, C.W. & Hungate, B.A. Increased soil emissions of potent greenhouse gases under increased atmospheric CO₂. *Nature* **475**, 214-216 (2011).
154. Lamsen, E.N. & Atsumi, S. Recent progress in synthetic biology for microbial production of C₃-C₁₀ alcohols. *Frontiers in Microbiology* **3** (2012).
155. Keasling, J.D. & Chou, H. Metabolic engineering delivers next-generation biofuels. *Nat Biotech* **26**, 298-299 (2008).

156. Okamoto, S., Ikeuchi, M. & Ohmori, M. Experimental analysis of recently transposed insertion sequences in the cyanobacterium *Synechocystis* sp. PCC 6803. *DNA research : an international journal for rapid publication of reports on genes and genomes* **6**, 265-273 (1999).
157. Barnes, W.M. PCR amplification of up to 35 kb DNA with high fidelity and high yield from bacteriophage templates. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2216-2220 (1994).
158. Mohamed, A. & Jansson, C. Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant Molecular Biology* **13**, 693-700 (1989).
159. Mohamed, A., Eriksson, J., Osiewacz, H.D. & Jansson, C. Differential expression of the *psbA* genes in the cyanobacterium *Synechocystis* 6803. *Molecular & General Genetics* **238**, 161-168 (1993).
160. You, L. et al. Metabolic pathway confirmation and discovery through ¹³C-labeling of proteinogenic amino acids. *J Vis Exp*, e3583 (2012).
161. Feng, X. et al. Mixotrophic and photoheterotrophic metabolism in *Cyanothece* sp. ATCC 51142 under continuous light. *Microbiology* **156**, 2566-2574 (2010).
162. Wahl, S.A., Dauner, M. & Wiechert, W. New tools for mass isotopomer data evaluation in (13)C flux analysis: mass isotope correction, data consistency checking, and precursor relationships. *Biotechnol Bioeng* **85**, 259-268 (2004).
163. Huang, H.-H., Camsund, D., Lindblad, P. & Heidorn, T. Design and characterization of molecular tools for a Synthetic Biology approach towards developing cyanobacterial biotechnology. *Nucleic acids research* **38**, 2577-2593 (2010).
164. Grosjean, D. Atmospheric chemistry of alcohols. *Journal of the Brazilian Chemical Society* **8**, 433-442 (1997).
165. Tichy, M. & Vermaas, W. In Vivo Role of Catalase-Peroxidase in *Synechocystis* sp. Strain PCC 6803. *Journal of Bacteriology* **181**, 1875-1882 (1999).
166. Cameron, J.C. & Pakrasi, H.B. Glutathione Facilitates Antibiotic Resistance and Photosystem I Stability during Exposure to Gentamicin in Cyanobacteria. *Applied and Environmental Microbiology* **77**, 3547-3550 (2011).
167. Shi, Z.P., Zhang, C.Y., Chen, J.X. & Mao, Z.G. Performance evaluation of acetone-butanol continuous flash extractive fermentation process. *Bioprocess and Biosystems Engineering* **27**, 175-183 (2005).
168. Kraemer, K., Harwardt, A., Bronneberg, R. & Marquardt, W. Separation of butanol from acetone-butanol-ethanol fermentation by a hybrid extraction-distillation process. *Computers & Chemical Engineering* **35**, 949-963 (2010).
169. Green, E.M. Fermentative production of butanol - the industrial perspective. *Current Opinion in Biotechnology* **22**, 337-343 (2011).
170. Ezeji, T., Milne, C., Price, N.D. & Blaschek, H.P. Achievements and perspectives to overcome the poor solvent resistance in acetone and butanol-producing microorganisms. *Applied Microbiology and Biotechnology* **85**, 1697-1712 (2010).
171. Zhang, C.C., Jeanjean, R. & Joset, F. Obligate phototrophy in cyanobacteria: more than a lack of sugar transport. *Fems Microbiology Letters* **161**, 285-292 (1998).

172. Liu, X.Y., Sheng, J. & Curtiss, R. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6899-6904 (2011).
173. Deng, M.D. & Coleman, J.R. Ethanol synthesis by genetic engineering in cyanobacteria. *Applied and Environmental Microbiology* **65**, 523-528 (1999).
174. Gao, Q., Wang, W., Zhao, H. & Lu, X. Effects of fatty acid activation on photosynthetic production of fatty acid-based biofuels in *Synechocystis* sp. PCC 6803. *Biotechnology for Biofuels* **5** (2012).
175. Lan, E.I. & Liao, J.C. ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6018-6023 (2012).
176. Frost, J.W. et al. in Carbon Management: Implications for R & D in the Chemical Sciences and Technology (A Workshop Report to the Chemical Sciences Roundtable) (The National Academies Press, Washington, D.C.; 2001).
177. Piringir, O.G. & Baner, A.L., Edn. 2nd (Wiley-VCH, Weinheim; 2008).
178. Vijayakumar, J., Aravindan, R. & Viruthagiri, T. Recent trends in the production, purification and application of lactic acid. *Chemical and Biochemical Engineering Quarterly* **22**, 245-264 (2008).
179. Garlotta, D. A literature review of poly(lactic acid). *Journal of Polymers and the Environment* **9**, 63-84 (2001).
180. Taskila, S. & Ojamo, H. in Lactic Acid Bacteria - R & D for Food, Health and Livestock Purposes. (ed. M. Kongo) (InTech, Rijeka, Croatia; 2013).
181. Wang, B., Pugh, S., Nielsen, D.R., Zhang, W. & Meldrum, D.R. Engineering cyanobacteria for photosynthetic production of 3-hydroxybutyrate directly from CO₂. *Metabolic engineering* **16**, 68-77 (2013).
182. Kusakabe, T. et al. Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metabolic engineering* **20C**, 101-108 (2013).
183. Wang, B., Wang, J., Zhang, W. & Meldrum, D.R. Application of synthetic biology in cyanobacteria and algae. *Front Microbiol* **3**, 344 (2012).
184. Berla, B.M. et al. Synthetic biology of cyanobacteria: unique challenges and opportunities. *Front Microbiol* **4**, 246 (2013).
185. Wu, G.F., Shen, Z.Y. & Wu, Q.Y. Modification of carbon partitioning to enhance PHB production in *Synechocystis* sp PCC6803. *Enzyme and Microbial Technology* **30**, 710-715 (2002).
186. Huang, H.H. & Lindblad, P. Wide-dynamic-range promoters engineered for cyanobacteria. *Journal of Biological Engineering* **7**, 10 (2013).
187. Joseph, A. et al. Utilization of Lactic Acid Bacterial Genes in *Synechocystis* sp. PCC 6803 in the Production of Lactic Acid. *Bioscience, Biotechnology, and Biochemistry* **77**, 966-970 (2013).
188. Angermayr, S.A., Paszota, M. & Hellingwerf, K.J. Engineering a Cyanobacterial Cell Factory for Production of Lactic Acid. *Applied and Environmental Microbiology* **78**, 7098-7106 (2012).

189. Angermayr, S.A. & Hellingwerf, K.J. On the Use of Metabolic Control Analysis in the Optimization of Cyanobacterial Biosolar Cell Factories. *The Journal of Physical Chemistry B* **117**, 11169-11175 (2013).
190. Wang, Q., Ingram, L.O. & Shanmugam, K.T. Evolution of D-lactate dehydrogenase activity from glycerol dehydrogenase and its utility for D-lactate production from lignocellulose. *Proceedings of the National Academy of Sciences* **108**, 18920-18925 (2011).
191. McNeely, K., Xu, Y., Bennette, N., Bryant, D.A. & Dismukes, G.C. Redirecting Reductant Flux into Hydrogen Production via Metabolic Engineering of Fermentative Carbon Metabolism in a Cyanobacterium. *Applied and Environmental Microbiology* **76**, 5032-5038 (2010).
192. Campbell, W.H. & Gowri, G. Codon Usage in Higher-Plants, Green-Algae, and Cyanobacteria. *Plant Physiology* **92**, 1-11 (1990).
193. Lindberg, P., Park, S. & Melis, A. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic engineering* **12**, 70-79 (2010).
194. Ungerer, J. et al. Sustained photosynthetic conversion of CO₂ to ethylene in recombinant cyanobacterium *Synechocystis* 6803. *Energy & Environmental Science* **5**, 8998-9006 (2012).
195. Vermaas, W.F.J. in eLS (John Wiley & Sons, Ltd, 2001).
196. Tamoi, M., Miyazaki, T., Fukamizo, T. & Shigeoka, S. The Calvin cycle in cyanobacteria is regulated by CP12 via the NAD(H)/NADP(H) ratio under light/dark conditions. *The Plant Journal* **42**, 504-513 (2005).
197. Takahashi, H., Uchimiya, H. & Hihara, Y. Difference in metabolite levels between photoautotrophic and photomixotrophic cultures of *Synechocystis* sp PCC 6803 examined by capillary electrophoresis electrospray ionization mass spectrometry. *Journal of Experimental Botany* **59**, 3009-3018 (2008).
198. Wermuth, B. & Kaplan, N.O. Pyridine nucleotide transhydrogenase from *Pseudomonas aeruginosa*: Purification by affinity chromatography and physicochemical properties. *Archives of Biochemistry and Biophysics* **176**, 136-143 (1976).
199. Bricker, T.M. et al. The Malic Enzyme Is Required for Optimal Photoautotrophic Growth of *Synechocystis* sp. Strain PCC 6803 under Continuous Light but Not under a Diurnal Light Regimen. *Journal of Bacteriology* **186**, 8144-8148 (2004).
200. Wendisch, V.F., de Graaf, A.A., Sahm, H. & Eikmanns, B.J. Quantitative Determination of Metabolic Fluxes during Coultivation of Two Carbon Sources: Comparative Analyses with *Corynebacterium glutamicum* during Growth on Acetate and/or Glucose. *Journal of Bacteriology* **182**, 3088-3096 (2000).
201. Heidorn, T. et al. Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods Enzymol* **497**, 539-579 (2011).
202. Niederholtmeyer, H., Wolfstadter, B.T., Savage, D.F., Silver, P.A. & Way, J.C. Engineering cyanobacteria to synthesize and export hydrophilic products. *Applied and Environmental Microbiology* **76**, 3462-3466 (2010).
203. Gao, Z., Zhao, H., Li, Z., Tan, X. & Lu, X. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy & Environmental Science* **5**, 9857-9865 (2012).

204. Gründel, M., Scheunemann, R., Lockau, W. & Zilliges, Y. Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology* **158**, 3032-3043 (2012).
205. Antoniewicz, M.R., Kelleher, J.K. & Stephanopoulos, G. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. *Analytical Chemistry* **79**, 7554-7559 (2007).
206. Tang, Y. et al. Central metabolism in *Mycobacterium smegmatis* during the transition from O₂-rich to O₂-poor conditions as studied by isotopomer-assisted metabolite analysis. *Biotechnology Letters* **31**, 1233-1240 (2009).
207. Brosius, J., Erfle, M. & Storella, J. Spacing of the -10 and -35 regions in the tac promoter. Effect on its in vivo activity. *J Biol Chem* **260**, 3539-3541 (1985).
208. Axe, D.D. & Bailey, J.E. Transport of lactate and acetate through the energized cytoplasmic membrane of *Escherichia coli*. *Biotechnol Bioeng* **47**, 8-19 (1995).
209. Jones, D.T. & Woods, D.R. Acetone-butanol fermentation revisited. *Microbiol Rev* **50**, 484-524 (1986).
210. Atsumi, S. et al. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metabolic engineering* **10**, 305-311 (2008).
211. Baez, A., Cho, K.-M. & Liao, J. High-flux isobutanol production using engineered *Escherichia coli*: a bioreactor study with in situ product removal. *Applied Microbiology and Biotechnology*, 1-10 (2011).
212. Luli, G.W. & Strohl, W.R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl Environ Microbiol* **56**, 1004-1011 (1990).
213. Tang, Y.J., Meadows, A.L. & Keasling, J.D. A kinetic model describing *Shewanella oneidensis* MR-1 growth, substrate consumption, and product secretion. *Biotechnol Bioeng* **96**, 125-133 (2007).
214. Nielsen, D.R. et al. Engineering alternative butanol production platforms in heterologous bacteria. *Metab Eng* **11**, 262-273 (2009).
215. Huo, Y.X. et al. Conversion of proteins into biofuels by engineering nitrogen flux. *Nature biotechnology* **29**, 346-351 (2011).
216. Trinh, C.T., Li, J., Blanch, H.W. & Clark, D.S. Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol. *Appl Environ Microbiol* **77**, 4894-4904 (2011).
217. Li, H. et al. Integrated electromicrobial conversion of CO₂ to higher alcohols. *Science* **335**, 1596 (2012).
218. Shen, C.R. & Liao, J.C. Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metabolic engineering* **10**, 312-320 (2008).
219. Lynd, L.R. et al. How biotech can transform biofuels. *Nat Biotech* **26**, 169-172 (2008).
220. Oncel, S. & Sabankay, M. Microalgal biohydrogen production considering light energy and mixing time as the two key features for scale-up. *Bioresour Technol* **121**, 228-234 (2012).
221. Zhang, F. et al. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic engineering* **14**, 653-660 (2012).

222. Zheng, Y.N. et al. Optimization of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered *Escherichia coli*. *Microbial Cell Factories* **11** (2012).
223. Passell, H. et al. Algae biodiesel life cycle assessment using current commercial data. *Journal of Environmental Management* **129**, 103-111 (2013).
224. Paddon, C.J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-532 (2013).
225. Keasling, J.D. Synthetic biology and the development of tools for metabolic engineering. *Metabolic engineering* **14**, 189-195 (2012).
226. Keasling, J.D. Manufacturing Molecules Through Metabolic Engineering. *Science* **330**, 1355-1358 (2010).
227. Yim, H. et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nature chemical biology* **7**, 445-452 (2011).
228. Lee, J.W. et al. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nature chemical biology* **8**, 536-546 (2012).
229. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nature biotechnology* **27**, 946-950 (2009).
230. Dueber, J.E. et al. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature biotechnology* **27**, 753-759 (2009).
231. Moon, T.S., Dueber, J.E., Shiue, E. & Prather, K.L. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metabolic engineering* **12**, 298-305 (2010).
232. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
233. Warner, J.R., Reeder, P.J., Karimpour-Fard, A., Woodruff, L.B.A. & Gill, R.T. Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat Biotech* **28**, 856-862 (2010).

Appendix 1: Microbial metabolisms and cell culture models for biofuel production. Bioenergy: Principles and Applications.

CHAPTER 13

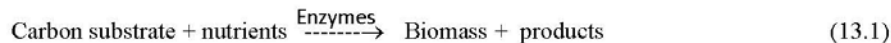
Microbial metabolisms and cell culture models for biofuel production

What is included in this chapter?

This chapter covers microbial metabolism as it is related to biofuel production. Discussions on microbial (both kinetic and metabolic) models to predict cell growth, the dynamics of product secretion, metabolite turnover rates, and yield coefficients are included. Pertinent examples and calculations have also been provided.

13.1 Introduction

Microbial metabolism involves complex sets of biochemical reactions catalyzed by different enzymes. These enzyme-mediated reactions are responsible for the synthesis of a large number of metabolites and generation of cellular energy. A microbial medium must be supplemented with carbon substrates and other essential nutrients (water, oxygen, nitrogen, phosphorus, etc.,) for all these reactions to take place. The progress of all the biochemical reactions within a cell is reflected as the production of more microbial mass and other products.



As such, a microbe operates like a very complex cell factory in which thousands of reactions take place. Microbial metabolisms can produce diverse bioproducts from cheap substrates, and thus microbial hosts are widely used for synthesizing biofuel molecules. In recent years, molecule biology technologies have been employed extensively to improve cell metabolic capability for converting renewable carbon compounds such as starch, cellulose and carbon dioxide into biofuels. Thereby, the understanding of the common metabolic pathways for biofuel conversion is important for development of advanced biofuel platforms. Moreover, metabolic models play a key role in scaling up of bioprocesses from laboratory to industrial level and make production economically viable. For example, microbial growth models quantify and predict biomass and product syntheses under various cultivation conditions (note: Monod-based model is commonly used to describe the rate of cell growth as the function of substrate concentrations). Other models such as stoichiometry flux analysis and elemental balance models measure intracellular reaction rates and the carbon flows from substrates to biomass and final products. They identify bottleneck reactions within the metabolic network, calculate the product and biomass yields, and predict the performance of the cell towards product synthesis. In this chapter, we introduce several useful microbial models, which can be used to describe microbial metabolisms and bioreactor operations for biofuel production.

The Authors are: **Arul Varman, Lian He, and Yinjie J. Tang**, School of Engineering & Applied Science, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, USA.

13.2 Carbon metabolisms in biofuel production

Microbial metabolisms can be divided into three categories based on the types of carbon sources:

Autotrophs–Inorganic carbon such as CO₂ is the carbon source. Plants, algae and cyanobacteria can harvest light energy and obtain electrons from water to produce biofuel precursor (such as lipids) and H₂. They use the Calvin cycle for CO₂ fixation, which employs the enzyme ribulose biphosphate carboxylase/oxygenase (RuBisCO) to convert ribulose-1,5-biphosphate and CO₂ into 3-phosphoglycerate (3PG). Some bacterial species may employ other autotrophic pathways, such as reductive citric acid cycle, Wood-Ljungdahl pathway, and 3-hydroxypropionate cycle.

Heterotrophs–Organic carbons are the carbon source. Glucose is one of the most common carbon sources employed by heterotrophs for biofuel production. Glucose catabolism occurs through three primary pathways: (i) Glycolysis (i.e., the Embden Meyerhof Parnas pathway); (ii) the pentose phosphate pathway; (iii) the Entner-Doudoroff pathway (ED). Microbial species can also use other carbon substrates (acetate, glycerol and xylose) for biomass growth and product synthesis (Figure 13.1).

Mixotrophs– CO₂ and organic carbon substrates are consumed simultaneously. For example, algal species show higher biomass and biofuel productivity in mixotrophic conditions.

The primary pathways of microbial metabolism are listed in Table 13.1. Microbes use these pathways to generate energy in the form of ATP (Adenosine-5'-triphosphate) and reducing equivalents such as NADH (Nicotinamide adenine dinucleotide), NADPH (Nicotinamide adenine dinucleotide phosphate), and FADH₂ (Flavin adenine dinucleotide). These pathways also provide chemical precursors that are essential to synthesize biofuel products (e.g., ethanol, propanol, and butanol) and biomass building blocks (e.g., amino acids, DNA - Deoxyribonucleic acid, RNA – Ribonucleic acid, lipids, and carbohydrates).

Table 13.1 Primary pathways for substrate metabolism

Pathway	Starting metabolite	Ending metabolite	ATP and reducing equivalents	Biomass precursors
Glycolysis	Glucose	2 Pyruvate	2 ATP, 2 NADH	3PG, Pyruvate, PEP
Oxidative PP pathway	G6P	C5P and CO ₂	2 NADPH	C5P, E4P
ED pathway	6PG	Pyruvate and GAP	none	3PG, Pyruvate
TCA cycle	Pyruvate	3 CO ₂	6 NADH, 2 FADH ₂ , 2 GTP	Acetyl-CoA, α-ketoglutarate, Oxaloacetate
Calvin cycle	3 CO ₂	GAP	-6 NADPH, -9 ATP*	3PG

* : “-” means consumption of energy molecules

Sugars (such as glucose, xylose starch, and sucrose) have been widely used for biofuel production, which can be obtained either from food crops (corn, sugarcane, sugar beet) or from biomass polymers (i.e., cellulose and hemicellulose). Non-sugar-based substrates, such as glycerol, lactate, acetate, CO₂, and syngas (CO, CO₂ and H₂), can also be converted into biofuels (Figure 13.1). Substrate yield coefficients are the ratio of the amount of product or biomass formed to the amount of substrate consumed. The theoretical yield dictates the maximum amount of a biofuel that can be produced from a given carbon source as tabulated in Table 13.2.

Ethanol, currently the most commercially successful biofuel, can be produced by yeast fermentations. Yeast efficiently converts sugar into ethanol and CO₂ via glycolysis pathway and pyruvate decarboxylase / alcohol dehydrogenase. High quality biofuels with properties similar to those of gasoline and diesel fuel are being synthesized by microorganisms (Figure 13.1). Several engineered biofuel pathways are being examined. For example, engineered *Escherichia coli* can use the keto acid pathway and the Ehrlich pathway to produce higher alcohols (such as isobutanol), while the mevalonate pathway in yeast can be extended to synthesize branched and cyclic hydrocarbons (the biofuels with lower freezing point and higher energy content).

Table 13.2 Theoretical biofuel yields of selected carbon substrates

Substrate	Product	Product yield (gram/gram)	Biomass yield (gram/gram)
Glucose	Ethanol	0.51	0.51
Glucose	Isobutanol	0.41	--
Glucose	lipid	0.32	--
Acetate	Fatty acids	0.29	0.36
Acetate	CH ₄	0.27	--
Glycerol	Ethanol	0.50	0.50
CO ₂ (with H ₂)	Methane	0.36	0.62

Finally, microbial metabolisms for biofuel production are very different across the species. *Saccharomyces cerevisiae* and *Escherichia coli* are microbial cell factories that are widely used in biofuel industrial because the two model species ferment sugar efficiently and are also amenable to genetic modification and bioprocess scale up. Other microbial species, such as cyanobacteria, are also promising hosts for biofuel production because they can convert sunlight and CO₂ to biomass and products. The species diversity in metabolic features offers opportunity for synthesizing many different useful products from diverse carbon substrates. Table 13.3 shows several different microbial species that produce biofuels, either via the native biofuel pathway or via a metabolically engineered pathway.

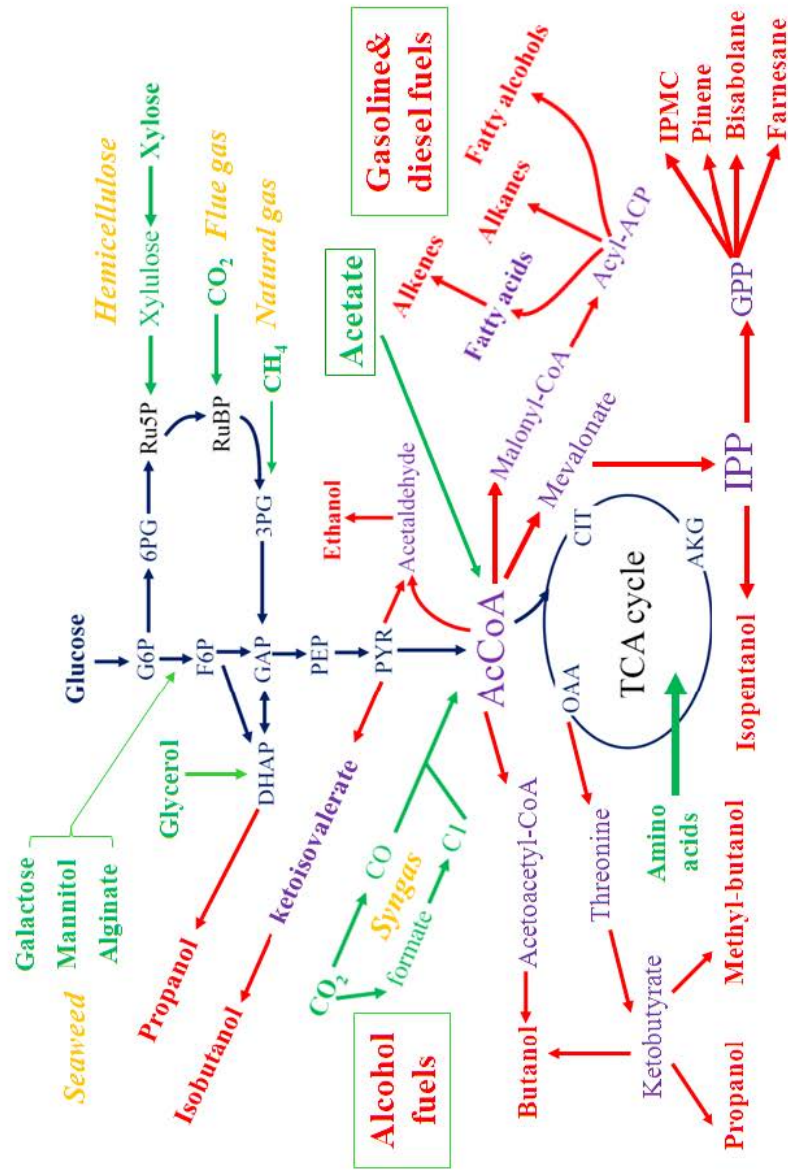


Figure 13.1 Summary of microbial metabolisms for producing different biofuels from diverse carbon substrates

Table 13.3. Commonly employed microbes for biofuel production

Species	Substrates	Products	Features
<i>Saccharomyces cerevisiae</i>	Glucose, fructose, galactose, and others	Alcohols	Easy genetic manipulations, Crabtree effect
<i>Zymomonas mobilis</i>	Glucose, fructose, sucrose	Ethanol	High ethanol tolerance and yield
<i>Clostridium thermocellum</i>	Glucose, cellulose, cellobiose	Ethanol	Growth at high temperature, mixed fermentation pathways
<i>Clostridium acetobutylicum</i>	Glucose, xylose	Ethanol and butanol	Acetone, ethanol, and butanol fermentation
<i>Escherichia coli</i>	Glucose, xylose, glycerol, and others	Alcohols, diesels, and other biofuels	Easy genetic manipulations, fast growth
Cyanobacteria (e.g., <i>Synechocystis 6803</i>)	CO ₂	Alcohols, H ₂ , fatty acids	CO ₂ fixation
<i>Phanerochaete chrysosporium</i>	Glucose and lignin	cellulosicbiomass pretreatment	Strong ability to degrade lignin
<i>Yarrowia lipolytica</i>	Glucose, acetate and fatty acids	Lipids	Oleaginous yeast that accumulates lipids

Crabtree effect: This is a phenomenon wherein, yeast is able to produce ethanol under oxygen conditions when provided with high concentrations of glucose.

13.3 Metabolic models for biofuel production

13.3.1 Microbial growth in batch culture

In batch system, microbial growth takes place in a closed system without inflow or outflow except aeration. Microbial growth is commonly monitored by measuring the dry weight, turbidity (optical density) of the culture medium, or the number of colony forming units (CFUs). For slow growing microbes (at very low cell density) or microbial culture containing suspended solids, cell density can be indirectly determined by measuring the concentrations of DNA, RNA, ATP, or total proteins.

Microbial growth usually undergoes five growth phases: (1) lag phase, (2) exponential phase, (3) declining rate of growth, (4) stationary phase, and (5) death phase (Figure 13.2). However, depending on the cultivation conditions, microbial growth curves may not strictly follow the five growth phases.

Lag Phase: In the lag phase, cell numbers remain unchanged for a while after the initial inoculation. The cells are metabolically active during this period and the cell size may increase. The lag phase can be reduced by using an inoculum from an exponential phase, by increasing the inoculation ratio, or by improving the nutrient conditions in the culture medium.

Exponential Growth Phase: Exponential growth phase (also known as logarithmic growth phase) starts when cells divide at a constant rate. In this stage, the metabolic activity (metabolic flux) and chemical composition of all the cells can be assumed to be in a pseudo-steady state. The exponential growth rate can be mathematically described as:

$$\frac{dX}{dt} = \mu X \quad (13.2)$$

where X (g/L) is the cell concentration and μ (hr^{-1}) is the specific growth rate for the cells. Integration of equation (13.2) from ($t = t_0, X = X_0$) to any other time point in the exponential growth phase (t, X) yields:

$$\ln\left(\frac{X}{X_0}\right) = \mu(t - t_0) \quad (13.3)$$

Equation (13.3) indicates a semi log plot of $\ln(X)$ versus time is a straight line with a slope of μ . The time required for a cell to double its population (t_d) is given by:

$$t_d = \frac{\ln(2)}{\mu} = \frac{0.693}{\mu} \quad (13.4)$$

A modified form of Eq. (13.2) includes the cell growth with a lag phase (t_L):

6

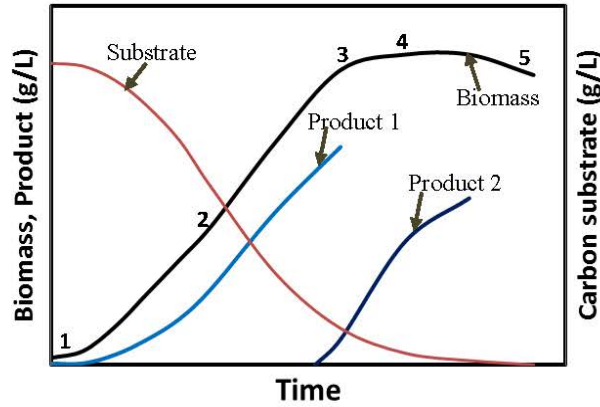


Figure 13.2 Biomass growth and product secretion curves. Product 1 is growth associated, while product 2 is non-growth associated.

$$\frac{dX}{dt} = \mu X \left(1 - e^{-t/t_l} \right) \quad (13.5)$$

Declining Growth Phase: At the end of exponential phase, one or more of the nutrients in the cell culture medium becomes limited for biomass growth, while the product accumulation may also stress cell metabolisms.

Stationary Phase: During this phase, the cell population remains constant because nutrients become limited and toxic products become inhibitory. At this stage, microbial host may actively produce some non-growth-associated products (such as antibiotics and lipids). Cells may continue to grow slowly, which is counterbalanced by cell death as discussed in the following phase.

Death Phase: During the death phase, cell lysis occurs and the cell population starts declining. The death phase can be represented by a first order rate equation of the form

$$\frac{dX}{dt} = -k_d X \quad (13.6)$$

where k_d is the first order decay rate constant.

Example 13.1:

Optical density (OD) measurement is normally employed as a direct means for monitoring the growth of microbes because OD values are proportional to cell concentrations. Growth of a blue green algal strain was monitored by measuring the absorbance of the culture at 730 nm and the data recorded is as given below:

Time (hrs):	0	21	96	128
OD ₇₃₀	0.4	0.6	3.2	?

Questions: Calculate the maximum specific growth rate and doubling time of this algal strain. Estimate the OD at time point 168 hrs.

Solution

$$\text{Specific growth rate } (\mu) = \frac{\ln(OD_2) - \ln(OD_1)}{t_2 - t_1} = \frac{\ln(3.2) - \ln(0.6)}{96 - 21} = 0.022 \text{ hr}^{-1}$$

$$\text{Doubling time} = \frac{\ln(2)}{\mu} = \frac{0.693}{0.022} = 31.5 \text{ hrs.}$$

$$OD_{168} = OD_{128} \exp [\mu (t_3 - t_2)] = 3.2 \exp [0.022(168 - 96)] = 6.5$$

13.3.2 Monod equation and inhibition kinetics

Cell growth can be described by the Monod equation. The Monod equation has been developed from observations on the growth of cells at different substrate concentrations and can fit a wide range of data. The kinetics described by this equation is similar to those of the Michaelis-Menten equation developed for enzymatic reactions. The Monod equation for cell growth limited by a single substrate is described by:

$$\mu = \frac{\mu_{\max} S}{K_s + S} \quad (13.7)$$

where μ_{\max} is the maximum specific growth rate. K_s is known as the half velocity constant or the Monod constant, and is equal to the substrate concentration at which the growth rate is half the maximum. The values of μ_{\max} and K_s depends upon the organism selected, type of substrate used, and cultivation conditions. The model takes two simpler forms depending on the substrate concentration:

At high substrate concentration ($S \gg K_s$),

$$S/(S+K_s) \approx 1, \text{ and } \mu = \mu_{\max} \text{ (0}^{\text{th}} \text{ order)} \quad (13.8)$$

At low substrate concentration ($S \ll K_s$),

$$S/(S+K_s) \approx S/K_s,$$

$$\text{and } \mu = \frac{\mu_{\max}}{K_s} S \text{ (1}^{\text{st}} \text{ order)} \quad (13.9)$$

As discussed above, at low substrate concentration (13.9), the Monod equation represents first order kinetics, that is, the growth rate is proportional to the substrate concentration. At high substrate concentration (13.8), the Monod equation approaches 0th order, that is, the growth rate is independent of substrate concentration. The Monod equation is graphically shown in Figure 13.3.

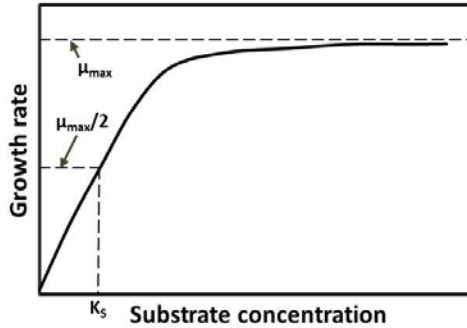


Figure 13.3 Microbial growth rates as a function of substrate concentrations

A Monod model that accounts for cell death with a biomass decay rate constant k_d (hr^{-1}), is:

$$\mu = \frac{\mu_{\max} S}{K_s + S} - k_d \quad (13.10)$$

The Monod model treats a cell population as a single homogenous system. However, there are other more complicated growth models. For example, segregated models treat each cell individually and view a cell culture as a heterogeneous mixture of cell population. Structured models represent a multicomponent approach that considers components within a cell, such as the concentration of intracellular metabolites, DNA, and RNA. These complex models have

been well described elsewhere and beyond the scope of the book (Blanch and Clark, 1997; Shuler and Kargi, 2000; Stephanopoulos, Aristidou and Nielsen, 1998).

13.3.3 Inhibition models

Substrates and products in the culture medium can inhibit cell growth. For example, high concentration of sugar decreases cell growth, while products such as ethanol and butanol interfere with cell physiology. A common approach to model inhibition accounts for the inhibitors (I) as competitive, uncompetitive or noncompetitive:

For a competitive inhibition model;

$$\mu = \frac{\mu_{\max} S}{K_s \left(1 + \frac{I}{K_i}\right) + S} \quad (13.11)$$

For an uncompetitive inhibition model;

$$\mu = \frac{\mu_{\max} S}{K_s + S \left(1 + \frac{I}{K_i}\right)} \quad (13.12)$$

For a noncompetitive inhibition model;

$$\mu = \frac{\mu_{\max} S}{K_s \left(1 + \frac{S}{K_s}\right) \left(1 + \frac{I}{K_i}\right)} \quad (13.13)$$

Another common inhibitory model is expressed by using the inhibitory constant in an exponential form;

$$\mu = \frac{\mu_{\max} S}{[K_s + S]} e^{-K_i I} \quad (13.14)$$

By using the maximum inhibitor concentration (I_m), we can get:

$$\mu = \frac{\mu_{\max} S}{[K_s + S]} \left(1 - \frac{I}{I_m}\right)^n \quad (13.15)$$

where, n is a parameter that indicates the toxicity of the inhibitor. As the value of I approaches I_m , the specific growth rate (μ) of cell becomes zero, indicating that the cell is no longer viable.

Multiple substrate models: The model for multi-substrates ($S_1, S_2, S_3, \dots, S_n$) can be written either as an additive or a multiplicative form. Eq. 13.16 describes cell growth by consuming several carbon substrates additively. For example, microbes use both glucose and xylose for biomass growth. Eq. 13.17 describes multiple-nutrient-controlled cell growth. For example, aerobic microbial growth is controlled by two substrates in a multiplicative manner (S_1 : glucose and S_2 : oxygen).

$$\mu = \frac{\mu_{\max 1} S_1}{[K_1 + S_1]} + \frac{\mu_{\max 2} S_2}{[K_2 + S_2]} + \dots + \frac{\mu_{\max n} S_n}{[K_n + S_n]} \quad (13.16)$$

$$\mu = \mu_{\max} \left[\frac{S_1}{[K_1 + S_1]} \frac{S_2}{[K_2 + S_2]} \dots \frac{S_n}{[K_n + S_n]} \right] \quad (13.17)$$

13.3.4 Monod-based kinetic model for biofuel production in batch bioreactors

Kinetic models can be used to describe biomass growth, substrate utilization, and product formation as a function of time. To link the biomass growth with substrate consumption and product formation, yield coefficients are used in a kinetic model. Yield coefficients are the ratios of amount of biomass or products produced to the amount of substrates consumed. To demonstrate how to simulate biofuel fermentations, a kinetic model to predict aerobic production of isobutanol by an engineered *Escherichia coli* is discussed in this section.

In a batch reactor, engineered *E.coli* can use glucose as the only carbon substrate to produce isobutanol (IB) and acetate. Isobutanol formation is both growth and non-growth associated as it is produced from an intermediate of an amino acid synthesizing pathway, which is active under all the growth stages of a microbial culture (Figure 13.2). Along with isobutanol, the strain excretes acetate as a growth associated fermentation product. Thus, the model consists of four time-dependent variables: X, A, IB, and G, which represent the concentrations of biomass, acetate, isobutanol, and glucose, respectively.

$$\text{Biomass: } \frac{dX}{dt} = R_X - k_d X \quad (13.18)$$

$$\text{Acetate: } \frac{dA}{dt} = R_A \quad (13.19)$$

$$\text{Isobutanol formation: } \frac{dIB}{dt} = R_{IB} \beta X - k_{IB} IB \quad (13.20)$$

$$\text{Glucose consumption: } \frac{dG}{dt} = -\frac{R_X}{Y_{XG}} - \frac{R_A}{Y_{AG}} - \frac{R_{IB}}{Y_{IBG}} \quad (13.21)$$

Biomass growth equation:
$$R_x = \frac{\mu_{max} G}{(K_s + G) \left(1 + \frac{A}{k_{iA}}\right)} X \quad (13.22)$$

Acetate production equation: $R_A = \alpha_{AX} R_x$ and (13.23)

Isobutanol production equation: $R_{IB} = \alpha_{IBX} R_x$ (13.24)

In the above equations, β is the non-growth associated isobutanol production rate; k_d is the cell death rate constant; and Y_{XG} , Y_{AG} , and Y_{IBG} are the yield coefficients for the production of biomass, acetate, and isobutanol from glucose, respectively. k_{IB} represents the rate of IB loss from the fermentor (due to product vaporization). The presence of acetate inhibits cell growth, and hence, a non-competitive inhibition constant K_{iA} is included in the model. α_{AX} and α_{IBX} are the growth-associated yield coefficients of acetate and IB, respectively.

Example 13.2:

For the parameter values given below, predict the isobutanol concentration as a function of time for 50 hours. Assume the initial concentrations of biomass, acetate, isobutanol, and glucose to be 0.12 g/L, 0.08 g/L, 0 g/L, and 20 g/L, respectively.

Parameters	Units	Values
K_s , Monod constant	g/L	0.305
μ_{max} , Maximum specific growth rate	1/h	0.073
K_{iA} , Acetate inhibition constant	g/L	47.69
k_d , Death rate constant	1/h	0.007
Y_{XG} , Yield coefficient for biomass	g biomass/g glu	0.161
Y_{IBG} , Yield coefficient for IB	g IB/g glu	0.163
Y_{AG} , Yield coefficient for acetate	g acetate/g glu	0.080
α_{IBX} , Growth associated yield of IB	g IB/g biomass	0.528
α_{AX} , Growth associated yield of acetate	g acetate/g biomass	0.614
k_{IB} , IB loss rate	1/h	0.123
β , Non-growth associated IB production rate	g IB / (g biomass·h)	0.006

Solution

The equations (13.18-13.24) are solved using the function “ode23” from MATLAB. The results are presented in Figure 13.4. Detailed MATLAB code for this problem can be found in the appendix (Appendix 13.1).

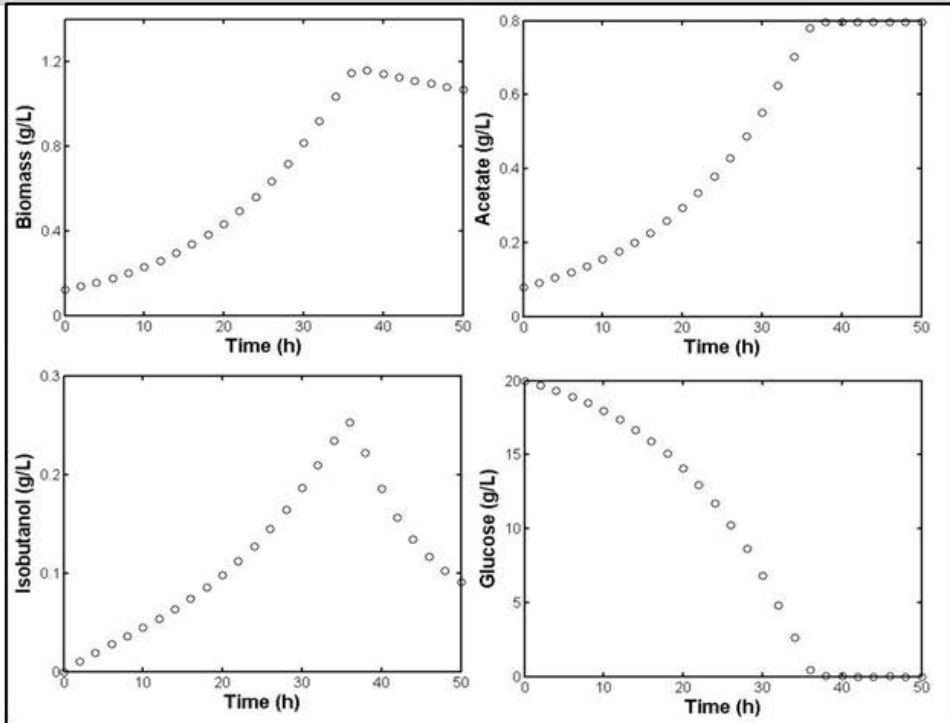


Figure 13.4 Simulated results for the isobutanol kinetic model, showing the cell growth, acetate production, isobutanol production, and glucose consumption with time.

Discussion: Figure 13.4 shows that isobutanol production is high at the early growth stage. When glucose is used up, the cell growth approaches its stationary phase and the isobutanol concentration starts declining (lost through bioreactor outlet under aeration conditions).

13.3.5 Monod model coupled with mass transfer

The presence of a liquid film around a cell can alter the substrate concentration to which the cell is exposed to. Hence, the substrate concentration at the cell surface, S_s , will be less than the substrate concentration in the bulk, S_o . This difference in substrate concentration will affect the cell growth rate and will be dependent on the mass transfer of the substrate from the bulk solution to the cell surface. The rate of substrate consumption for cell growth is:

$$\text{Rate of substrate consumption} = \frac{\mu_{\max} S_s X}{[K_s + S_s] Y_{X/S}} \quad (13.25)$$

The substrate transfer from the bulk solution to the cell surface can be described by a mass transfer coefficient k_L (1/hr), interfacial area per unit volume of the cell a (1/m), and the ratio of cell concentration to cell density (X/ρ):

$$\text{Rate of substrate transport} = k_L a \left(\frac{X}{\rho} \right) (S_0 - S_s) \quad (13.26)$$

At steady state, there is no accumulation of substrate within the liquid film, and the rate of substrate transfer to the cell surface will be equal to the rate at which substrate is consumed. Hence, by equating (13.25) and (13.26), we get

$$\frac{\mu_{\max} S_s X}{[K_s + S_s] Y_{X/S}} = k_L a \left(\frac{X}{\rho} \right) (S_0 - S_s) \quad (13.27)$$

After rearranging and simplifying equation (13.27), we get

$$\mu = \frac{\mu_{\max} S_0}{S_0 + K_s + \frac{\mu_{\max} \rho}{k_L a Y_{X/S}}} \quad (13.28)$$

The detailed steps and examples involved in equation (13.28) can be found in Blanch and Clark (1997). Equation (13.28) clearly reveals that rate of growth is dependent on both mass transfer coefficient and size of the cell.

13.3.6 Mass balances and reactions in fed batch and continuous-stirred tank bioreactor

In a fed batch bioreactor, fresh media is added to the bioreactor without product removal. A fed batch bioreactor is commonly employed to produce biofuel products. In a continuous bioreactor (i.e., chemostat), fresh media is continuously added into the bioreactor, and at the same time the culture medium containing products, wastes, and cells is continuously removed from the bioreactor. Figure 13.5 shows a continuous flow bioreactor with S_{in} and S_{out} representing the concentrations of substrate in the inlet and outlet. Q_{in} and Q_{out} are volumetric flow rates at the inlet and outlet, respectively. V is volume of the reactor, r_s is substrate consumption rate by cells, and X is the biomass concentration.

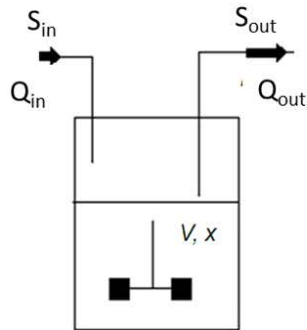


Figure 13.5 Continuous flow bioreactor

A general mass balance on substrate is given by:

$$\frac{dS}{dt} = -r_s X + \frac{Q_{in}}{V} S_{in} - \frac{Q_{out}}{V} S_{out} - S \frac{1}{V} \frac{dV}{dt} \quad (13.29)$$

where V , S and X represent the volume of the bioreactor, the concentrations of substrate and biomass in the bioreactor respectively.

In a **continuous-stirred tank bioreactor**, $\frac{dS_{out}}{dt} = \frac{dV}{dt} = 0$,

$Q_{in} = Q_{out} = Q$, and $D = \frac{Q}{V}$ where D is the dilution rate.

Substituting these values in equation (13.29), we get:

$$r_s X = D(S_{in} - S_{out}) \quad (13.30)$$

In a **fed-batch bioreactor**, $Q_{out} = 0$, and by substituting this in equation (13.29), we get:

$$\frac{dS}{dt} = -r_s X + \frac{Q_{in}}{V} S_{in} - S \frac{1}{V} \frac{dV}{dt} \quad (13.31)$$

Further simplification, by substituting $\frac{dV}{dt} = Q_{in}$ along with the definition of dilution rate in equation (13.31) yields:

$$\frac{dS}{dt} + r_s X = D(S_{in} - S) \quad (13.32)$$

Moreover, mass balance expressions for biomass and product formations in fed batch and continuous-stirred bioreactors can be derived using similar approach (equation 13.29), which are described elsewhere (Stephanopoulos, Aristidou and Nielsen, 1998).

Example 13.3:

“Washout” is a condition in which the dilution rate ‘D’ is too high and the cells cannot grow fast enough leaving the biomass concentration in the reactor to be $X = 0$. For the parameter values, $\mu_{\max} = 0.2 \text{ hr}^{-1}$, $K_S = 1 \text{ g/L}$, and $S_0 = 10 \text{ g/L}$, calculate the maximum dilution rate above which washout will occur. Find the flow rate within which a 1 liter bioreactor must be operated without washout.

Solution

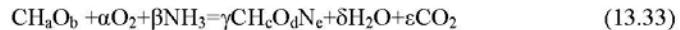
At steady state, $\mu = D$. Under washout condition $X = 0$ and henceforth the substrate concentration inside the bioreactor $S = S_0$.

$$\text{Maximum dilution rate, } D_{\max} = \frac{\mu_{\max} S_0}{[K_S + S_0]} = \frac{(0.2)(10)}{(10 + 1)} = 0.18 \text{ hr}^{-1}.$$

$$\text{Maximum flow rate, } Q_{\max} = (V) (D_{\max}) = 0.18 \text{ L/hr.}$$

13.3.7 Elemental balance and stoichiometric models**Elementary balance model**

The approximate chemical formula of microbial biomass is given by $\text{CH}_{1.8}\text{O}_{0.6}\text{N}_{0.2}\text{S}_{0.004}\text{P}_{0.01}$. Carbon thus constitutes nearly 50% of dry cell weight of biomass; and is an important element required for the growth of the cell. Since the mass of each element is conserved, elementary balance model can be used to determine the stoichiometry of a biological reaction when the compositions of substrates and products are known. The following reaction represents the general aerobic growth of a microorganism ($\text{CH}_c\text{O}_d\text{N}_e$) on an organic carbon substrate (CH_aO_b) and a nitrogen source (NH_3):



The respiratory quotient (RQ) can be measured under aerobic growth conditions, and is defined as the ratio of the CO_2 production rate to the O_2 consumption rate:

$$\text{RQ} = \frac{\epsilon}{\alpha} \quad (13.34)$$

The elementary balance equations can be written in a matrix form and the elementary balance of equation (13.33) can be calculated via a matrix algorithm.

$$\begin{array}{l} \text{Carbon} \\ \text{Oxygen} \\ \text{Nitrogen} \\ \text{Hydrogen} \\ \text{RQ} \end{array} \begin{pmatrix} 0 & 0 & \alpha & 1 & 1 & 0 & 1 \\ -2 & 0 & \beta & b & d & 1 & 2 \\ 0 & -3 & \gamma & a & c & 2 & 0 \\ 0 & -1 & \delta & 0 & e & 0 & 0 \\ -\text{RQ} & 0 & \epsilon & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} \quad (13.35)$$

Stoichiometric model

The stoichiometric model is for the estimation of metabolic fluxes in a denoted metabolic network. The basic equation is:

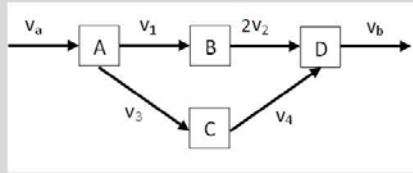
$$\frac{dc}{dt} = S \cdot v - b \quad (13.36)$$

where, dc/dt is the net accumulation rate of all metabolites, S is the stoichiometry matrix for a specific network, v is the vector for the flux variables, and b is the vector for the substrate uptake, biomass synthesis, and product formation rates. In the matrix S , each column contains the stoichiometry of a specific reaction in the network, and the row corresponds to mass balance of intermediates, substrates, or products. At steady-state condition, the inflow rates of all the metabolites are equal to their outflow rates. Hence, the equation can be simplified as:

$$S \cdot v = b \quad (13.37)$$

Example 13.4:

Consider the simple network shown below. We can calculate the internal fluxes v_1 , v_2 , v_3 and v_4 using the stoichiometric model, assuming that v_a and v_b are determined experimentally.



The fluxes for v_1 , v_2 , v_3 and v_4 can be calculated with a stoichiometric model as follows:

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} v_a \\ 0 \\ 0 \\ v_b \end{pmatrix}$$

Abbreviations

3PG – 3-Phosphoglycerate
6PG – 6- phosphogluconolactone
AcCoA – Acetyl CoA
AKG – α -ketoglutarate
ATP – Adenosine-5'-triphosphate
C1 – 5,10-Methylenetetrahydrofolate;

C5P – Ribose-5-phosphate
CIT – Citrate
DHAP – Dihydroxyacetone phosphate
E4P – erythrose-4-phosphate;
F6P – Fructose 6-phosphate
FADH₂ – Flavin adenine dinucleotide
GAP – Glyceraldehyde 3-phosphate
G6P – Glucose 6-phosphate
GPP – Geranyl pyrophosphate
GTP – Guanosine-5'-triphosphate
IPMC – 1-Isopropyl-4-methylcyclohexane
IPP – Isopentenyl pyrophosphate
NADH – Nicotinamide adenine dinucleotide
NADHP – Nicotinamide adenine dinucleotide phosphate
OAA – Oxaloacetate
PYR – Pyruvate
Ru5P – Ribulose-5-phosphate
RuBP – Ribulose-1,5-bisphosphate

References

- Bailey, J. E. 1998. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress* 14:8-20.
- Clark, D. S., and H. W. Blanch. 1997. *Biochemical Engineering*, Second Edition. Taylor & Francis.
- Dellomonaco, C., F. Fava, and R. Gonzalez. 2010. The path to next generation biofuels: successes and challenges in the era of synthetic biology. *Microbial Cell Factories* 9:3.
- Kauffman, Kenneth J., Purusharth Prakash, and Jeremy S. Edwards. 2003. Advances in flux balance analysis. *Current opinion in biotechnology* 14: 491-496.
- Lee, Jong Min, Erwin P. Gianchandani, and Jason A. Papin. Flux balance analysis in the era of metabolomics. *Briefings in Bioinformatics* 7: 140-150.
- Lin, J., S.-M. Lee, H.-J. Lee, and Y.-M. Koo. 2000. Modeling of typical microbial cell growth in batch culture. *Biotechnology and Bioprocess Engineering* 5:382-385.
- Orth, Jeffrey D., Ines Thiele, and Bernhard O. Palsson. 2010. What is flux balance analysis?. *Nature biotechnology* 28: 245-248.
- Pamela, P. P.-Y., Z. Fuzhong, B. d. C. Stephen, and D. K. Jay. 2012. Microbial engineering for the production of advanced biofuels. *Nature* 488:320-328.

Pelczar Jr, M. J., E. C. S. Chan, and N. R. Krieg. 1993. Microbiology. Tata McGraw-Hill.

Pramanik, J., and J. D. Keasling. 1997. Stoichiometric model of Escherichia coli metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. Biotechnology and Bioengineering 56:398-421.

Pratt, C. W., D. Voet, and J. G. Voet. 2008. Fundamentals of Biochemistry. Wiley John & Sons.

Shuler, M. L., and F. Kargi. 2002. Bioprocess Engineering: Basic Concepts. Prentice Hall.

Stephanopoulos, N., A. Aristidou, and J. Nielsen. 1998. Metabolic engineering principles and methodologies. Academic Press, San Diego.

Xiao, Y., X. Feng, A. M. Varman, L. He, H. Yu, and Y. J. Tang. 2012. Kinetic Modeling and Isotopic Investigation of Isobutanol Fermentation by Two Engineered Escherichia coli Strains. Industrial & Engineering Chemistry Research 51:15855-15863.

Exercise Problems

14.1. For inhibition models, which type of inhibition reduces the model's apparent μ_{\max} ?

14.2. *Zymomonas mobilis* uses the ED pathway for the fermentation of glucose. How many NADHs are produced when the cells convert glucose to pyruvate via the ED pathway?

14.3 Describe two different butanol producing pathways.

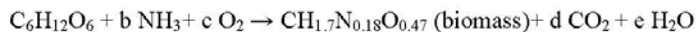
14.4. How many ATPs/NADHs are generated by glycolysis? Which pathway produces the majority of NADPH?

14.5. One microbial species can use either acetate or glycerol as the sole carbon source for growth. Can you predict which carbon substrate yields higher biomass?

14.6. What is the advantage of using *Clostridium thermocellum* to produce ethanol? Can yeast ferment ethanol under aerobic condition?

14.7. What is the model equation for uncompetitive substrate inhibition for microbial growth?

14.8. The biomass formation equation for a microorganism using glucose and NH_3 as carbon and nitrogen sources is shown below.



Balance the above equation and calculate the glucose yield coefficient for biomass ($c/d=1.9$).

14.9. An engineered *E.coli* has a mass doubling time of 3 h when grown on acetate. The Monod constant using acetate is 1 g/L, and cell yield on acetate is 0.4g cell/g acetate. If we operate a chemostat on a feed stream containing 40 g/L acetate, find the substrate concentration and cell productivity (g cell/h) when the chemostat dilution rate is 60 % of theoretical maximum dilution rate.

14.10. Bioreactions are often carried out in batch reactors. Using the available information to determine how much time is required to achieve a 95% conversion of the substrate. Assume that the volume V of the reactor content is constant, and the reaction rate follows Monod Model.

Initial conditions: Biomass $X(0)=0.05$ g/L, Substrate $S(0)=15$ g/L, Product $P(0)=0$ g/L

Parameter values: $V=1$ L, $\mu_{\max}=0.2$ hr⁻¹, $K_s=1$ g/L

Yield coefficients are assumed to be constant: $Y_{x/s}=0.5$ g biomass/g substrate, $Y_{p/s}=0.1$ g product/g substrate.

Appendix 13.1

Code used for example 13.1

```
% initial conditions
y0=[0.12 0.08 0 20]; %x A IB G
tspan=[0:2:50];

% solving the differential equations "ff"
[t1,y]=ode23(@ff,tspan,y0);
y1=y(:,1); %X
y2=y(:,2); %A
y3=y(:,3); %Ib
y4=y(:,4); %G

%Plot
figure(1)
plot(t1,y(:,1),'ko')
set(gca,'linewidth',1.5,'fontsize',12,'YTick',0:0.4:1.2)
xlabel('Time (h)','fontsize',16,'fontweight','b')
ylabel('Biomass (g/L)','fontsize',16,'fontweight','b')

figure(2)
plot(t1,y(:,2),'ko')
set(gca,'linewidth',1.5,'fontsize',12,'YTick',0:0.2:1)
xlabel('Time (h)','fontsize',16,'fontweight','b')
ylabel('Acetate (g/L)','fontsize',16,'fontweight','b')

figure(3)
plot(t1,y(:,3),'ko')
set(gca,'linewidth',1.5,'fontsize',12,'YTick',0:0.1:0.3)
ylim([0 0.3])
xlabel('Time (h)','fontsize',16,'fontweight','b')
ylabel('Isobutanol (g/L)','fontsize',16,'fontweight','b')

figure(4)
plot(t1,y(:,4),'ko')
set(gca,'linewidth',1.5,'fontsize',12,'YTick',0:5:20)
ylim([0 20])
xlabel('Time (h)','fontsize',16,'fontweight','b')
ylabel('Glucose (g/L)','fontsize',16,'fontweight','b')

function dy=ff(t,y)
    param=[0.073 0.305 47.69 0.007 0.614 0.528 0.123 0.161 0.080 0.163 0.006];
    Ug=(param(1)*y(4)/(param(2)+y(4)))*(1/(1+y(2)/param(3)));
    Rx=y(1)*Ug;
    dy(1)=y(1)*(Ug-param(4)); %biomass
    dy(2)=param(5)*Rx; % acetate
    dy(3)=param(6)*Rx+param(11)*y(1)-param(7)*y(3); % isobutanol
    dy(4)=-Rx/param(8)-dy(2)/param(9)-param(6)*Rx/param(10); % glucose
    dy=dy';
end
```

Appendix 2: Engineering *Escherichia coli* to convert acetic acid to free fatty acids.



Regular article

Engineering *Escherichia coli* to convert acetic acid to free fatty acidsYi Xiao^{a,*}, Zhenhua Ruan^b, Zhiguo Liu^b, Stephen G. Wu^a, Arul M. Varman^a, Yan Liu^b, Yinjie J. Tang^{a,**}^a Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO 63130, USA^b Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:

Received 1 February 2013

Received in revised form 5 April 2013

Accepted 12 April 2013

Available online xxx

Keywords:

¹³C

acs gene

Anaerobic-digested

Carbon flux

Lignocellulosic biomass

ABSTRACT

Fatty acids (FAs) are promising precursors of advanced biofuels. This study investigated conversion of acetic acid (HAc) to FAs by an engineered *Escherichia coli* strain. We combined established genetic engineering strategies including overexpression of *acs* and *tesA* genes, and knockout of *fadE* in *E. coli* BL21, resulting in the production of ~1 g/L FAs from acetic acid. The microbial conversion of HAc to FAs was achieved with ~20% of the theoretical yield. We cultured the engineered strain with HAc-rich liquid wastes, which yielded ~0.43 g/L FAs using waste streams from dilute acid hydrolysis of lignocellulosic biomass and ~0.17 g/L FAs using effluent from anaerobic-digested sewage sludge. ¹³C-isotopic experiments showed that the metabolism in our engineered strain had high carbon fluxes toward FAs synthesis and TCA cycle in a complex HAc medium. This proof-of-concept work demonstrates the possibility for coupling the waste treatment with the biosynthesis of advanced biofuel via genetically engineered microbial species.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Food based materials are the most widely used fermentation feedstock for biofuels and biochemicals production [1,2]. However, considering the global food shortage and increasing cost of agriculture, non-food based substrates have been studied as alternative feedstocks. First, syngas, generated from various inexpensive sources such as natural gas and inedible biomass, can be used by some acetogens for alcohols and acetic acid production [1–4]. Second, lignocellulosic biomass can be hydrolyzed via chemical or biological methods to C5 and C6 sugars for microbial fermentations [5,6]. Third, algal phototrophic process converts CO₂ into lipid and other compounds. Fourth, other cheap feedstocks have found a niche in various biofuels production processes, including glycerol (for fatty acids) [7], protein-rich materials (for higher alcohols) [8], and CO₂ (via electromicrobial conversion to higher alcohols) [9]. However, non-food based biofuel production still faces challenges

(such as high cost of pretreatment, poor solubility of substrates, and low productivity).

This study attempts to use acetic acid (HAc) as a feedstock for production of biofuel. HAc can be derived from various cheap sources (Fig. 1). (1) HAc can be generated from syngas via chemical [10,11] (e.g., methanol carbonylation is an efficient route for HAc production [12,13]) and microbial [3,4,14] catalyses (e.g., HAc is a major product during syngas fermentation). (2) Methane from natural gas or biogas can be converted to HAc [15]. (3) HAc is present in several common waste streams. For example, it is a byproduct from hydrolysis (under acid or alkali pretreatment [16]) or pyrolysis of lignocellulosic biomass [17,18]. HAc is also an intermediate from anaerobic digestion (AD) of organic wastes. AD requires the co-operation of microbial communities through hydrolysis, acidogenesis, acetogenesis, and methanogenesis. By inhibiting methanogens, AD can accumulate HAc [19–22]. Oleaginous yeast (*Cryptococcus curvatus*) has been reported to utilize waste HAc as a main carbon substrate for lipid accumulation [23–25] and some *Clostridium* species can use HAc and sugar simultaneously for alcohols/butyrate syntheses [26,27]. However, the scarcity of tools for genetic manipulation on such organisms prevents those becoming workhorses for further improvement by metabolic engineering. Therefore, we sought to engineer a well-characterized microorganism *E. coli* to generate FAs from HAc. FAs biosynthesis by *E. coli* has attracted extensive interests because FAs are important precursors of fatty esters, fatty alcohols, waxes, and alkanes [28–30]. Recent studies on biosynthesis of FAs are summarized in Table 1. HAc has energy content comparable to glucose in

Abbreviations: AccABCD, acetyl CoA carboxylase; Acs, acetyl CoA synthase; AckA, acetate kinase; FadD, acyl CoA synthetase; FadE, acyl CoA dehydrogenase; PoxB, pyruvate oxidase; Pta, phosphate acetyltransferase; TesA, acyl ACP thioesterase; FAs, free fatty acids; HAc, acetic acid; IPTG, isopropyl β-D-1-thiogalactopyranoside; OD, optical density; YE, yeast extract; AD, anaerobic digestion; AA, amino acids.

* Corresponding author. Tel.: +1 314 935 3457; fax: +1 314 935 7211.

** Corresponding author. Tel.: +1 314 935 3441; fax: +1 314 935 7211.

E-mail addresses: huazhongxy@gmail.com (Y. Xiao), yinjie.tang@seas.wustl.edu (Y.J. Tang).

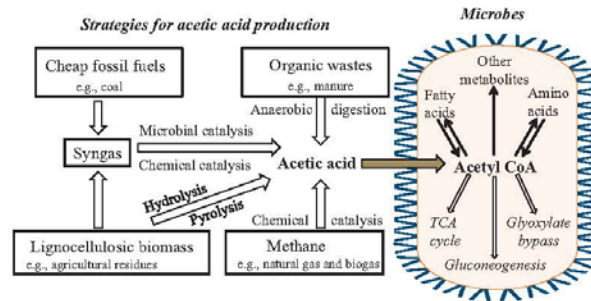


Fig. 1. Strategies for acetic acid production and its metabolism in microorganisms.

aerobic conditions (Fig. 2), but it is a notorious inhibitor for *E. coli* fermentation. It can diffuse into cell and interfere with cytoplasmic pH and henceforth enzymatic activities [31,32], thereby very few studies have employed genetically engineered microbes for conversion of HAC into advanced biofuels.

2. Materials and methods

2.1. Strains, plasmids, and culture conditions

Genetic manipulation was done according to the standard procedures [33] and gene deletion was performed using a commercial kit (Gene Bridges). All plasmids/strains and primers used in this study are listed in Tables 2 and S1, respectively. Plasmid constructions were described in Table S1.

M9 minimal medium contained 6.78 g/L Na_2HPO_4 , 3 g/L KH_2PO_4 , 0.5 g/L NaCl, 1 g/L NH_4Cl , 0.1 mM CaCl_2 (0.011 g/L), and 2 mM MgSO_4 (0.24 g/L). For HAC utilization experiments, unless stated otherwise, the engineered strains were incubated at 37 °C in shaking tubes (5 mL culture) at 200 rpm with the M9 medium containing different concentrations of sodium acetate, 0.2 g/L YE (yeast extract), 0.1 mM IPTG (isopropyl β -D-1-thiogalactopyranoside), 25 $\mu\text{g}/\text{mL}$ kanamycin, and 1x vitamin mix (Sigma).

Table 1
Summary of recent studies on fatty acids biosynthesis.

Substrates	Feature description	Titer (g/L)	Ref.
Glucose	Modular optimization of multi-gene pathways in <i>E. coli</i> and fed-batch fatty acids production	~8.6	[57]
Glucose	Reversal of the β -oxidation cycle; overexpression of thioesterase FadM from <i>E. coli</i>	~7	[58]
Glucose	Overexpression of thioesterase TesA and transcription factor FadR from <i>E. coli</i>	5.2	[59]
Glucose	Overexpression of thioesterase from plant and integrating the thioesterase gene into chromosome	~1	[60]
Glycerol	Overexpression of thioesterases and acetyl-CoA carboxylase	2.5	[7]
CO_2	A genetically modified cyanobacterium using CO_2 to produce and secrete fatty acids	~0.2	[61]

To screen engineered FAs producing strains, strains were initially grown at 37 °C in shaking tubes with 5x M9 medium containing 4.1 g/L sodium acetate (50 mM) and 5 g/L YE. When culture density reached $\text{OD}_{600} \sim 3$ (>16 h), 0.1–0.2 mM IPTG and another 50 mM sodium acetate were added and then the temperature was changed to 30 °C for FAs accumulation. The samples were collected at ~36 h (~20 h after IPTG addition). To optimize the cultivation for FAs production, 2.5 g/L HAC (equivalent final concentration) was added around ~24 h and ~30 h (i.e., 5 g/L HAC was resupplied before sample collection).

2.2. pH-coupled HAC fed-batch fermentations

To scale up FAs production from HAC and reduce the toxicity of HAC on cells, a pH-coupled fed-batch fermentation was performed [9]. The fermentation minimal medium contained 14.6 g/L K_2HPO_4 , 4 g/L KH_2PO_4 , 10 g/L $(\text{NH}_4)_2\text{SO}_4$, 2 g/L sodium citrate, 2.05 g/L sodium acetate, 0.117 g/L betaine, 0.011 g/L CaCl_2 , 0.72 g/L MgSO_4 .

Table 2
Plasmids and strains used in this study.

Plasmids/strains	Characteristics	Sources
<i>E. coli</i> strains		
DH5 α	For cloning	Invitrogen
BL21(DE3)	Wild type cell	Novagen
BL21(Δpta)	BL21(DE3) with deletion of <i>pta</i> gene	This study
BL21($\Delta poxB$)	BL21(DE3) with deletion of <i>poxB</i> gene	This study
BL21($\Delta pta/\Delta poxB$)	BL21(DE3) with deletion of <i>pta</i> and <i>poxB</i> genes	This study
BL21($\Delta fadE$)	BL21(DE3) with deletion of <i>fadE</i> gene	This study
Plasmids		
pUC19K	Kan ^r , <i>lac</i> promoter, a derivative of pUC19	PEER ^a
pTAC-MAT-Tag-2	Amp ^r , <i>tac</i> promoter, expression vector	Sigma
pMSD8	Plasmid carrying gene <i>accABCD</i> under a T7 promoter	Cronan Lab ^b
pYX30	pUC19K carrying <i>acs</i> gene	This study
pYX31	pUC19K carrying <i>ack</i> and <i>pta</i> genes	This study
pYX32	pUC19K carrying <i>acs</i> , <i>ack</i> , and <i>pta</i> genes	This study
pYX26	pTAC-MAT-Tag-2 carrying <i>tesA</i> gene	This study
pYX33	pYX30 carrying <i>tesA</i> gene under a <i>tac</i> promoter	This study

^a Power Environmental Energy Research Institute, California (provided by Dr. Qinhong Wang).

^b University of Illinois at Urbana-Champaign (provided by Dr. John Cronan).

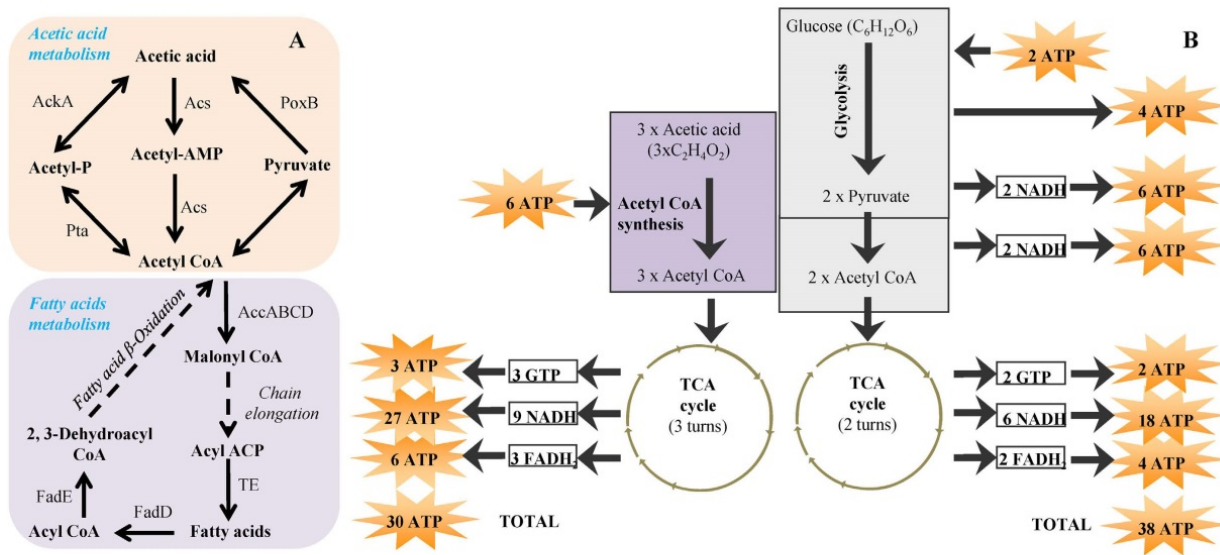


Fig. 2. The metabolic pathways of fatty acids biosynthesis from acetic acid (A) and the energy yield (B). The solid lines represent single step, while dotted lines represent multiple steps. Note: we assume that three ATP are generated from one NADH, while two ATP are generated from one FADH₂.

40 mg/L kanamycin, 1× vitamin (Sigma) and 1× trace metal solution. The 1000× trace metal solution (per liter) contained the following: $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$ 2.4 g, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 2.4 g, CoCl_2 0.26 g, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 0.48 g, $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ 0.24 g, H_3BO_3 3 g, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 8.4 g, Na_3EDTA 20.6 g. Fermentations were carried out in a 1 L New Brunswick Bioflo 110 system with a temperature electrode and a pH meter. In fermenter, 400 mL of culture was inoculated with 5 mL of overnight LB culture (BL21($\Delta fadE$)/pYX33). The pH was maintained at ~7.0 during the entire fermentation by adding pure HAC through an auto-pump. The stirring speed was kept at 250 rpm and the air rate was held at ~1 L/min. FAs fermentation using the minimal medium was very slow (the lag phase was over 50 h at 37 °C). To increase product yield and reduce the lag time, we tested different fermentation temperatures and nutrient conditions. The improved FAs fermentation was under a rich medium and a lower temperature (note: fermentation temperature can affect the stability and performance of the mutant [34,35]). Therefore, we used the rich medium (i.e., the minimal fermentation medium supplemented with 1% YE) for the fed batch culture and the fermentation temperature was at 25 °C. IPTG (0.2 mM) was added at 55.5 h and 13 h for fermentations using the minimal medium and the rich medium respectively.

2.3. Analytical methods

FAs measurement was based on previous reports [7,30,36,37]. A 175 μL sample of culture, spiked with 100 mg/L pentadecanoic acid (C15:0) as an internal standard, was extracted by 400 μL of methanol–chloroform (1:1) and acidified by ~5 μL of concentrated HCl. The organic layer, separated by centrifugation, was transferred to a new tube, and the solvent was removed under vacuum conditions. Methyl derivation of FAs was performed at 40 °C for ~2 h by adding 200 μL of the reagent mix (10 μL of concentrated HCl, 20 μL of (trimethylsilyl) diazomethane solution (2 M), and 170 μL of methanol). After that, the samples were vortexed with 200 μL of 0.9% NaCl and 350 μL of ethyl acetate, and then the organic layer was taken for GC–MS analysis. The methyl esters were analyzed using GC (Hewlett Packard 7890A, Agilent Technologies, equipped with a DB5-MS column, J&W Scientific) and a mass spectrometer (5975C, Agilent Technologies) under the following program: hold at 80 °C for 1 min, ramp to 280 °C at 30 °C/min⁻¹, and then hold at 280 °C for 3 min. The carrier gas was helium. The FAs methyl esters were quantified based on standard curves. Other than methyl ester of dodecenoic acid (C12:1) which was generated from derivatization of corresponding fatty acids, F.A.M.E. Mix(C8–C24), methyl oleate, methyl myristoleate, and methyl pentadecanoate were purchased from Sigma as standards.

Glucose, xylose, and HAC concentrations in the waste streams were measured using an HPLC (Shimadzu) equipped with a Bio-rad Aminex HPX-87H analytical column and a refractive index detector. The mobile phase was 5 mM sulfuric acid with a flow rate of 0.6 mL/min⁻¹, and column temperature was set at 65 °C [16]. The biomass was determined spectrophotometrically, where an OD₆₀₀ of one corresponded to 0.338 g dry weight/L for *E. coli* BL21 strains. The HAC was also measured via an enzyme kit (R-Biopharm). All the error bars presented in the results indicated the standard deviation from the mean ($n=2$ or 3).

2.4. ¹³C isotopic experiments for analyzing fatty acids yields

Isotopic tracer experiments were performed to provide fundamental information on the acetate utilization in the complex culture medium. We supplemented minimal medium with 0.25, 0.5, or 1% YE, and replaced the non-labeled HAC by fully labeled HAC (Cambridge Isotope Laboratories). The FAs production by ¹³C-cultures followed the same protocol described in Section 2.1. The

contributions of the YE and the HAC to the biomass and FAs were estimated by measuring the ¹³C-abundance in the metabolites. The isotopic analysis of 15 proteinogenic AAs (amino acids) was performed according to the previously described methods [38,39]. The ¹³C fractions (R_B , ¹³C carbons/total carbons) of AAs were calculated with the following equation:

$$R_B = \frac{1}{n} \sum_{x=0}^n (x \cdot M_x) \quad (1)$$

where n is the total carbon number of each AA, while x ($0 \leq x \leq n$) is the number of the labeled ¹³C in each AA. M_x is the corresponding ¹³C isotopomer fraction for each AA (M_0 is the fraction of the unlabeled fragment; M_1 is the fraction of the singly-labeled fragment, etc.).

Considering overlap of fragment ions, three methyl esters of fatty acids (C12:0, C14:1, and C14:0) were chosen, and the m/z [M]⁺ from the unfragmented methyl ester was used for ¹³C fraction analysis. R_F is the ratio of ¹³C labeled carbons to total carbons of the FAs, representing the contribution of ¹³C-HAC to FAs biosynthesis. We considered the natural abundance of ¹³C (1.1%) and one methyl group in the esters from the non-labeled methanol via methylation, thus the R_F were calculated using the Eq. (2):

$$R_F = \frac{n+1}{n} \left(\frac{100}{98.9} \frac{1}{n+1} \sum_{x=0}^{n+1} \left(x \frac{A_x}{\sum_{x=0}^{n+1} A_x} \right) - \frac{1.1}{98.9} \right) \quad (2)$$

where n is the total carbon number of each fatty acid, while x ($0 \leq x \leq n$) is the number of the labeled ¹³C in each fatty acid. A_x is the corresponding abundance of fragment ions in the mass spectrum (A_0, A_1, A_2, \dots are the corresponding abundances of the unlabeled, singly-labeled, doubly-labeled fragments, ...).

2.5. Conversion of acetic acid in waste streams to fatty acids

Preparing waste streams from hydrolysis of lignocellulosic biomass: Corn stover, switchgrass, and poplar were collected from fields at Michigan State University, while miscanthus and giant reed were obtained from Werks Management, LLC (Fishers, IN). Each feedstock was dried and ground using a mill. Dilute sulfuric acid pretreatment of these biomass samples was performed at 130 °C. Using 2% (w/w) acid, the switchgrass and miscanthus were treated for 2 h, while the corn stover was treated for 1 h. Using 3% (w/w) acid, the giant reed and the poplar were treated for 1 h and 2 h, respectively. The hydrolysates were collected and pH was neutralized to ~7 by addition of $\text{Ca}(\text{OH})_2$. The $\text{Ca}(\text{OH})_2$ could also precipitate and reduce fermentation inhibitors from the hydrolysates.

Preparing waste stream from anaerobic digestion: The sewage sludge for AD was from the East Lansing waste-water treatment plant. The sludge was condensed by centrifugation and a final 5% of total solid sludge was adopted for AD [40]. The AD was carried out using a 500 mL anaerobic bottle with 400 mL of the pretreated sludge medium (pH=7). At the beginning, 12.5% (v/v) of manure AD effluent was added as seed. Methanogens inhibitor (iodoform, 20 g/L) was added at 0.4 mL/L every other day, while pH was maintained using 30% (w/w) NaOH solution [41]. The AD effluent was harvested at 25 days by centrifuging at 4 °C. The effluent was autoclaved before using for *E. coli* culture.

Test of fatty acids production from waste streams: The mutant BL21($\Delta fadE$)/pYX33 was first grown in 3 mL of 5× M9 salt medium with 0.5% YE and 50 mM sodium acetate at 37 °C for ~16 h (OD₆₀₀ ~ 3). 1 mL of the HAC-rich waste streams and 0.2 mM IPTG were added into the pre-grown cultures ($t=0$ h), then the cultivation temperature was maintained at 30 °C for 24 h to accumulate FAs. The negative control used water to replace the waste streams.

3. Results and discussions

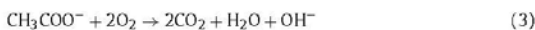
3.1. Optimization of HAC utilization in *E. coli*

HAC is a potential feedstock for synthesis of biomass and value-added chemicals. However, *E. coli* has a limited metabolic capability to assimilate HAC for biomass growth [42]. *E. coli* assimilates HAC via AMP-forming acetyl CoA synthetase (the *acs* pathway) and phosphotransacetylase/acetate kinase (the reversible PTA-ACKA pathway) [42]. To select an optimal HAC utilization route, both pathways were cloned, first separately and then together, into a vector pUC19K under a *lac* promoter, resulting in plasmids pYX30 (with the ACS pathway), pYX31 (with the PTA-ACKA pathway), and pYX32 (with both of the pathways) (Table 2). In acetate mediums (Fig. 3), BL21/pUC19K (containing an empty vector as a control), BL21/pYX31, and BL21/pYX32 could grow in 25 and 50 mM acetate medium, whereas their growths on 100 mM acetate were severely inhibited. The strain BL21/pYX30 showed the highest biomass growth and acetate tolerance (100 mM), whose *acs* pathway ($\text{ATP} + \text{acetate} + \text{CoA} \rightarrow \text{AMP} + \text{pyrophosphate} + \text{acetyl-CoA}$) produces AMP and requires additional ATP to convert AMP to ADP ($\text{AMP} + \text{ATP} \rightarrow 2\text{ADP}$).

On the other hand, the PTA-ACKA pathway consumes only one ATP per acetate for acetyl-CoA synthesis, but overexpression of this pathway impeded acetate utilization and biomass growth. The reverse PTA-ACKA pathway and the *PoxB* pathway can generate HAC from acetyl CoA and pyruvate respectively (Fig. 2A). We blocked the two competitive pathways to create mutant BL21(ΔpoxB , Δpta). Although the knockout strain had reduced HAC production in glucose-based cultures (Fig. S1A), the knockout strain bearing the plasmid pYX30 showed poor growth in acetate medium (Fig. S1B). Therefore, overexpressing the single *acs* gene along with maintaining the native HAC pathways was the best strategy for HAC assimilation in *E. coli* [34,43].

Degradation of HAC to CO_2 is the only energy production route for *E. coli* growth on HAC (Fig. 2B). When the medium contains

acetate salt, its degradation increases pH in the culture:



Therefore, we increased M9 salt concentrations to improve buffer capacity in our culture medium. For BL21/pYX30 culture, 5× M9 salt medium provided more stable acetate-based growth than normal M9 medium. Moreover, intermittent feeding with HAC during cell cultivation could adjust the medium's pH and re-supply carbon substrate to achieve higher biomass production.

3.2. Engineering an optimal fatty acids biosynthesis pathway

FAs biosynthesis [7,44] goes through multiple enzymatic steps (Fig. 2A). The reported metabolic strategies for FAs synthesis mainly include overexpression of thioesterase, enhancement of acetyl-CoA conversion to malonyl-CoA, and blocking fatty acids degradation through the β -oxidation pathway [7,30,45–47]. Recently, computational modeling successfully predicts gene targets through entire metabolic network for overproducing FAs in *E. coli* [48]. Although the effectiveness of some FAs synthesis strategies is still controversial among different literatures, acyl-ACP thioesterase is generally thought as the key enzyme for FAs production [7,30]. Therefore, a *tesA* gene from *E. coli* with leader sequence deleted, encoding cytosolic thioesterase, was cloned into the pTAC-MAT-Tag-2 expression vector (Sigma–Aldrich). The *tesA* gene with a *tac* promoter, a T1T2 terminator, and a repressor gene *lacI* was subcloned into pYX30, resulting in pYX33.

BL21/pYX33 produced 370 mg/L FAs from HAC medium with 0.5% YE, while BL21/pYX30 (negative control) produced less than 50 mg/L FAs (Fig. 4A). The composition of FAs included C8:0, C10:0, C12:1, C12:0, C14:1, C14:0, C16:1, C16:0, C18:1, and C18:0 (Fig. S2). Deleting acyl-CoA dehydrogenase (*fadE*) gene related to FAs degradation was reported to increase FAs production [30]. In our results, the BL21(ΔfadE)/pYX33 produced 20% more FAs than BL21/pYX33 (Fig. 4B). In addition, FAs were mainly located inside the cell and only 4% of the total FAs were found in the supernatant (Fig. 4C).

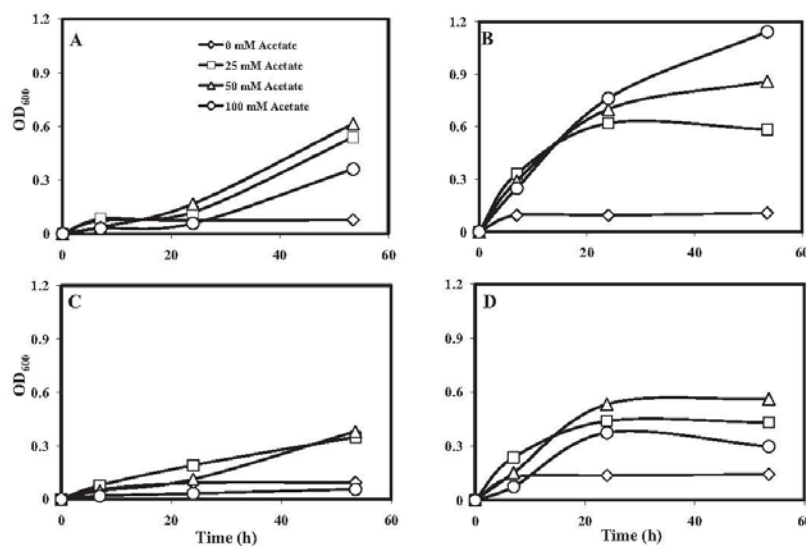


Fig. 3. Comparative growths of the engineered mutants using acetic acid. The mutants BL21/pUC19K (A), BL21/pYX30 (B), BL21/pYX31 (C), and BL21/pYX32 (D) grew on different concentrations of acetate. The OD was measured spectrophotometrically ($n=2$, the mean values were presented). The details of HAC utilization experiments are recorded in Section 2.

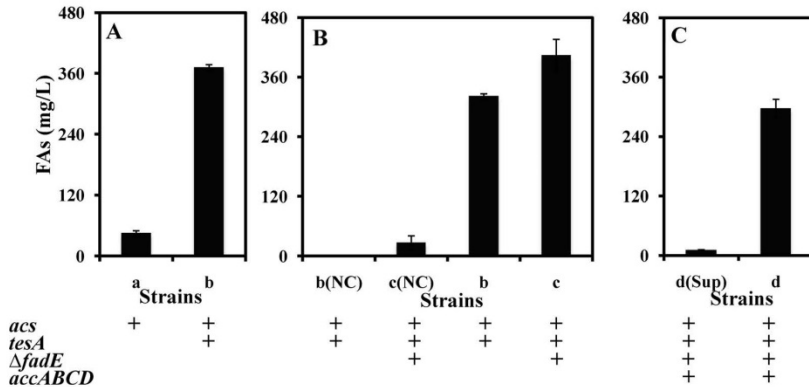


Fig. 4. Engineering *E. coli* strains for FAs production via HAC. (A) FAs productions by BL21/pYX33 and BL21(Δ *fadE*)/pYX33; (B) improvement of FAs production by deletion of gene *fadE*; (C) FAs productions by BL21(Δ *fadE*)/pYX33 and BL21(Δ *fadE*)/pYX33 + pMSD8. ($n=3$, culture medium with HAC and 0.5% YE). Abbreviations: (a) BL21/pYX33; (b) BL21(Δ *fadE*)/pYX33; (c) BL21(Δ *fadE*)/pYX33 + pMSD8; (d) BL21(Δ *fadE*)/pYX33 + pMSD8; NC, negative control (only 0.5% YE and 50 mM acetate were in the medium); Sup, the supernatant of culture.

Finally, the acetyl-CoA carboxylase (*AccABCD*) gene (borne on the pMSD8 [49] plasmid) was introduced into the BL21(Δ *fadE*)/pYX33, resulting in BL21(Δ *fadE*)/pYX33 + pMSD8. However, this mutant did not increase HAC-based biomass growth or FAs production. Although overexpressing the acetyl-CoA carboxylase has been suggested to improve FAs synthesis from glucose and glycerol [7,49], a combined in silico and experimental study points out that excess malonyl-CoA could be diverted toward formation of other components under certain growth conditions [48]. In this study, we think that the FAs precursor (malonyl-CoA) may not be a bottleneck for our strain to produce FAs from HAC.

3.3. Investigating the theoretical FAs yields

The stoichiometry of FAs synthesis is: 8 acetyl-CoA + 7ATP + 14NADPH \rightarrow fatty acid [50]. We assume that two ATP are required for producing one acetyl-CoA from HAC and NADPH is produced from NADH through transhydrogenases [51]: 8 HAC + 23 ATP + 14 NADPH (or NADH) \rightarrow fatty acid. If one mole of NADH is equivalent to three moles of ATP, one mole of fatty acid synthesis requires 65 moles of ATP that can be produced by oxidizing 6.5 moles of HAC (Fig. 2B). Thereby, one mole of FAs production consumes \sim 14.5 mol of HAC, resulting in the theoretical FAs yield as \sim 0.29 g FAs/g HAC (assuming FAs molecular weight as 256).

3.4. The effect of yeast extract on FAs production and yield

In previous biofuel reports, nutritional supplements, such as yeast extract, play an important role in improving fermentation performance [8,46,52–54]. In our experiments, the mutants (BL21/pYX33 and BL21(Δ *fadE*)/pYX33) produced less than 0.05 g/L FAs if 5 g/L YE was used as the sole carbon source. In the presence of YE and HAC, both biomass growth and FAs were promoted significantly (Table 3). To identify the contribution of YE to biomass and FAs production, we cultured BL21(Δ *fadE*)/pYX33 with fully 13 C-labeled HAC and non-labeled YE. Fifteen proteinogenic amino acids (AAs) were measured. Before addition of IPTG, the YE (12 C carbons) contributed over 75% AAs to biomass growth (Fig. 5A, most AAs were non-labeled). After 20 h IPTG induction, the 13 C abundance in most AAs increased by \sim 2 folds (Fig. 5B), which confirmed utilization of HAC for biomass synthesis.

Among these AAs, aspartate and glutamate (derived from the TCA cycle metabolites) showed the highest labeling abundance

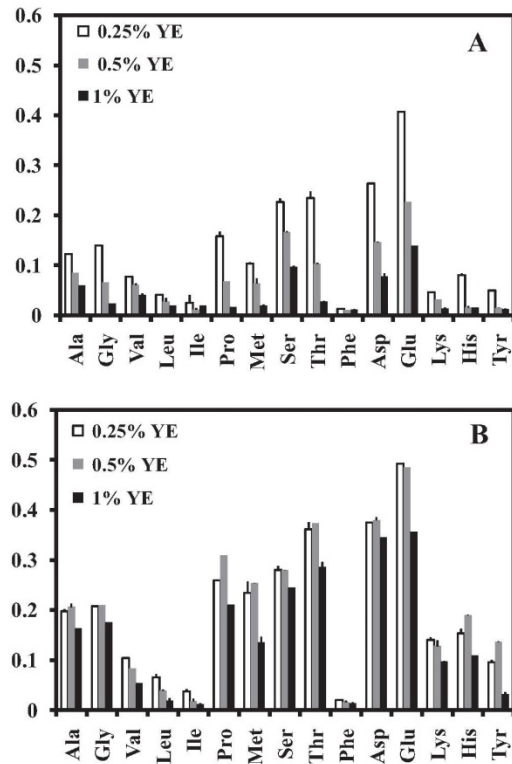


Fig. 5. 13 C-fractions in the proteinogenic amino acids from the engineered *E. coli*. The BL21(Δ *fadE*)/pYX33 grew on the medium with fully labeled 13 C-acetic acid and three different concentrations of non-labeled 0.25, 0.5, or 1% YE. The cells were harvested at 0h (the time point of IPTG addition) (A) and 20h (B), respectively. ($n=2$).

Table 3

Analysis of fatty acids biosynthesis via isotopic experiments. The BL21(Δ fadE)/pYX33 mutant was cultured in the 5 \times M9 salt medium containing 13 C labeled acetic acid (HAc) and three different concentrations (0.25, 0.5, or 1%) of non-labeled yeast extract (YE) in shaking tubes. The samples were taken at 0 h (the time at IPTG addition) and 20 h, respectively. Three single fatty acids (C12:0, C14:0, and C14:1) were selected for analysis ($n=2$).

Products (mg/L)	0.25% YE	0.5% YE	1% YE
Biomass (20h)	1440 \pm 10	2560 \pm 10	3080 \pm 10
FAs (0 h)	47.0 \pm 0.7	50.2 \pm 1.8	68.8 \pm 0.3
FAs (20 h)	433.1 \pm 5.9	570.7 \pm 0.2	999.7 \pm 2.2
13 C ratio - %			
FAs-C14:0 (0h)	40.3 \pm 1.5	39.1 \pm 6.2	35.6 \pm 2.3
FAs-C14:0 (20 h)	66.1 \pm 0.6	64.1 \pm 0.1	59.8 \pm 0.2
FAs-C14:1 (20 h)	76.6 \pm 1.3	68.4 \pm 1.8	65.3 \pm 0.9
FAs-C12:0 (20 h)	70.8 \pm 0.7	69.2 \pm 0.6	64.2 \pm 0.6

(up to ~40%), implying that HAc was mainly metabolized via the TCA cycle. Acetate based metabolism requires gluconeogenesis and pentose phosphate pathways (PPP) to synthesize precursors of several amino acids (histidine, phenylalanine and tyrosine). In our engineered strain, the labeling abundance in these amino acids was below 20%, indicating limited fluxes from acetate into gluconeogenesis and PPP pathways. Besides, phenylalanine (not tyrosine) was mostly non-labeled through the entire cultivation in rich medium, indicating that the engineered strain had very low capability for de novo biosynthesis of phenylalanine. Moreover, labeling results indicated that *E. coli* had a poor capability to synthesize leucine and isoleucine (both amino acids require pyruvate as the precursors) from HAc. This observation can be explained that metabolic flux from acetate to biomass was also limited at the pyruvate node (i.e., pyruvate synthesis was bottleneck step for biomass synthesis from HAc).

Isotopomers in FAs were analyzed after the cells grew with YE and 13 C labeled HAc (Fig. S3). Three FAs (C12:0, C14:1, and C14:0) were selected for quantifying HAc contribution to FAs production (Table 3). At the IPTG addition point (0 h), ~0.05 g/L FAs had been produced and ~60% of FAs were from YE. After 20 h, ~10 g/L HAc had been consumed and FAs reached 0.43, 0.57, and 1.0 g/L in the culture with 0.25, 0.5, and 1% YE respectively, in which the corresponding 13 C-abundance of FAs were 71 \pm 5%, 67 \pm 3%, and 63 \pm 3%. This observation confirmed that the key metabolic node (acetyl-CoA, precursor of FAs) in the engineered strain was mainly derived from acetate even in a complex culture medium. In theory, the maximum FAs yield from HAc is 0.29 g FAs/g HAc. Isotopic analysis allows us to determine the true yield of FAs from HAc in our strain. Given that 63% of the 1 g/L FAs were produced from HAc in the rich medium, the actual yield was ~0.063 g FAs/g HAc and equivalent to ~22% of the theoretical yield.

3.5. pH-coupled HAc fed-batch fermentation

To scale up FAs production from HAc, a pH-coupled HAc fed batch fermentation, which can gradually add HAc to the culture, was carried out in a 1-L bioreactor. When the fermentation used the rich medium with YE at room temperature, the biomass and FAs increased with no obvious lag phase and reached their maximum concentrations of 3.3 g/L and 0.9 g/L at 26 h, respectively (Fig. 6A). In the period from 13 h to 26 h, ~14 g/L HAc was consumed and 0.9 g/L FAs was generated. A FAs-producing mutant using glycerol (a major component in crude glycerol waste) was reported to have a high productivity (170 mg/L/h) and yield (48 mg FAs/g glycerol) [7]. Given that 65% of the FAs generated by the mutant were from HAc, the actual productivity was 45 mg/L/h and the actual yield was ~42 mg FAs/g HAc (~15% of the theoretical yield). On the other hand, the engineered strain had a poor growth in the complete minimal medium at 37 °C. There was a long lag phase (50 h) and the final FAs were only 0.45 g/L (~13% of the theoretical yield) after 100 h (Fig. 6B). The minimal acetate medium and higher growth temperature may cause more stresses on cells and reduce final biosynthesis productivity [34,35].

Another observation was that the compositions of the FAs from the cells grown in the rich medium and the minimal medium were rather different (Fig. S2). Within the rich medium (HAc + YE), the products of medium-chain FAs (C8–C14, desirable precursors for jet fuel, detergent or pharmaceuticals) were 68% of the total FAs, while the ratio of unsaturated FAs reached ~37%. The medium-chain FAs content from our engineered strain was notably higher than those from oleaginous yeast *C. curvatus* [25]. When the cells grew in the minimal medium, the long-chain FAs (C16–C18) comprised 56% of the total FAs, and the ratio of unsaturated FAs was decreased to 26%. In previous studies, the higher proportion of the medium-chain FAs has also been found in the LB rich medium compared with that in the defined minimal medium [7]. In silico analysis indicates that fatty acids chain length in *E. coli* is impacted by many central genes, transcriptional and post-transcriptional regulations, and FAs precursor pool sizes (e.g., acetyl-CoA) [48]. Thereby, *E. coli* FAs composition might be sensitive to the carbon substrates and growth conditions.

3.6. FAs production from waste HAc

HAc can be cheaply obtained from several types of waste streams. Although acetate can be chemically converted into other products, the compositions of waste streams are complicated with relative low concentration of acetate (<6%). Thereby, it is more economic to use biological method to convert waste acetate to other

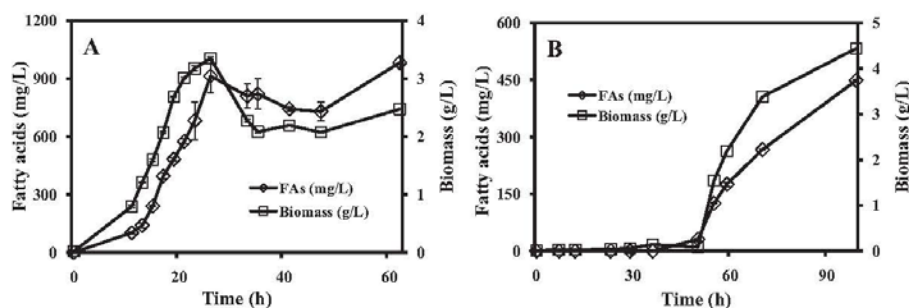


Fig. 6. pH-coupled acetic acid feeding fermentation. The fatty acids-producing mutant BL21(Δ fadE)/pYX33 grew in the rich medium (A) and the minimal medium (B). The FAs and biomass were measured ($n=2$).

value-added products. In this study, we applied the engineered *E. coli* for treating two acetate rich waste streams.

Acid pretreatment is an important step for lignocellulosic biofuel production. Hydrolysis of five different types of biomass materials (corn stover, switchgrass, poplar, miscanthus, and giant reed) by dilute acid generated waste streams containing xylose, HAc, and a small amount of furfurals and glucose. In such waste streams, it is still economically infeasible to recover these carbon sources. Therefore, we directly used these waste streams as feedstock for BL21(Δ *fadE*)/pYX33 fermentation. To prove this concept, the *E. coli* culture was first grown in the medium with acetate and YE for 16 h, and then the cultures were supplemented with IPTG inducer and each waste stream, using a mix ratio of 3 (the culture) to 1 (the waste stream solution). In these cultures (0h), glucose varied from 0.4 g/L to 1.3 g/L, xylose ranged from 2.2 g/L to 5.2 g/L, and HAc was from 3.3 g/L to 4 g/L (Fig. 7). After 24 h of IPTG induction, the mutants in each test had consumed all of the glucose, part of the xylose, and most of the HAc, generating up to 0.43 g/L FAs. The negative control (replacing the stream with water) contained only 0.035 g/L FAs, while the positive control (using xylose as sole carbon source) produced 0.24 g/L FAs. These results indicated that the engineered strain utilized both sugars and acetate in waste streams for FAs production.

AD process can accumulate significant acetate under specific conditions (up to 5% in the AD effluents) [55]. Here, we tested an AD effluent from municipal sludge (containing ~4 g/L HAc). The experiment was performed by adding 1 unit volume of waste effluent (along with IPTG) into 3 unit volumes of BL21(Δ *fadE*)/pYX33 culture ($OD_{600} > 3$, pre-grown in the acetate rich medium without IPTG induction). In the final cultures (24 h), we found that the cells had consumed all HAc (and possibly other unidentified nutrients from AD effluent) and generated 0.17 g/L FAs. For the control experiment, BL21(Δ *fadE*)/pYX33 culture mixed with water (3:1)

instead of AD effluent, the final culture only contained <0.04 g/L FAs. We recently compared an engineered *E. coli* strain modified with similar strategies to a wild-type fungus [56]. We found that both could use acetate to generate FAs (or lipids) and the engineered *E. coli* (with overexpressing a regulatory transcription factor *FadR*) had higher performance than the fungus.

3.7. Significance and future work

The use of acetate as a fermentation feedstock has advantages. (1) Sources for generating acetate are diverse and plentiful. (2) Acetate is miscible with water, allowing easy mass transfer during fermentation. (3) Acetyl-CoA, directly synthesized from acetate, is a starting precursor for many value-added compounds. As a proof of concept, we demonstrated that 1 m³ AD effluent (based on ~4 kg/m³ HAc) and 1 m³ waste streams from the dilute acid treatment of lignocellulosic biomass can produce about 0.54 and 0.8–1.6 kg FAs, respectively.

On the other hand, this study showed low performance of the engineered strain for acetate based biofuel production. This result implied that the host cell faced more stress in the presence of HAc, compared with other carbon sources (i.e., glucose or glycerol). Different metabolic engineering strategies are going to be tested in the future, including enhancement of *E. coli* acetate-tolerance, addition of quorum sensing system for auto induction of FAs production, design of dynamic sensor-regulator in the FAs pathway. Moreover, for the economic production of biofuels from acetate, we need to replace expensive yeast extract by other cheap protein-rich materials (e.g., corn steep liquor, algae residues or fermentation wastes [8]). Economic and life cycle analyses should also be performed to determine the applicability of this technology for both waste management and bioenergy applications.

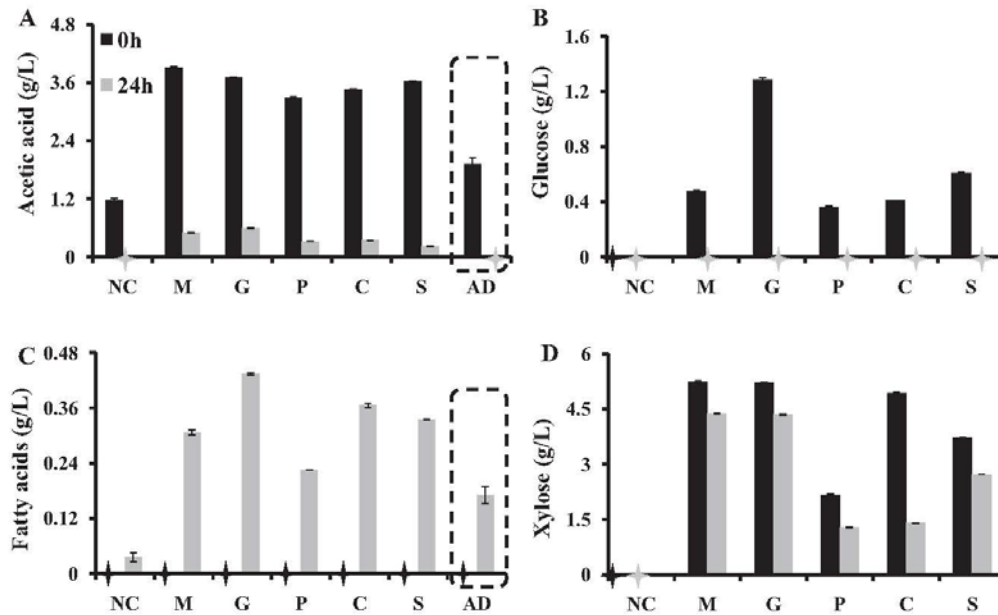


Fig. 7. Production of FAs using HAC-rich wastes. The time of IPTG addition was defined as 0 h. The acetic acid (A), glucose (B), fatty acids (C), and xylose (D) were measured ($n = 2$) at 0 h (black) and 24 h (gray). The star indicates the value is negligible. Samples: NC, water (negative control); AD, effluent from anaerobic digestion of sludge (boxed region); C, pretreated corn stover stream; S, pretreated switchgrass stream; P, pretreated poplar stream; M, pretreated miscanthus stream; G, pretreated giant reed stream.

4. Conclusions

This work overexpressed the HAC assimilation pathway and the FAs biosynthesis pathway in an *E. coli* strain. We demonstrate that the engineered strain can utilize acetic acid to produce 0.9 g/L FAs (within 26 h) and 0.45 g/L (within 100 h) in the rich and minimal medium respectively. Since acetate is inhibitory to *E. coli* growth, FAs titer obtained in this work was still much lower than sugar or glycerol based FAs production. However, this study opens opportunities to employ *E. coli* or any other species amenable for genetic engineering (such as fungal species and *Corynebacterium*) to economically utilize low concentration HAC as well as other nutrient sources in waste streams to achieve both environmental benefits and bioenergy harvesting.

Acknowledgements

We would like to thank Prof. John Cronan from University of Illinois at Urbana-Champaign for providing the plasmid pMSD8 and Prof. James Ballard for proofreading this paper. We also thank Hui Feng Yu, Ryan Lago, Amelia Chen, and Chang-Ki Kang for their help. The authors thank Bill & Melinda Gates Foundation for initiating this study and an NSF grant for ¹³C-metabolism analysis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bej.2013.04.013>.

References

- [1] C.R. Fischer, D. Klein-Marcuschamer, G. Stephanopoulos, Selection and optimization of microbial hosts for biofuels production, *Metab. Eng.* 10 (2008) 295–304.
- [2] D. Peters, Raw materials, *Adv. Biochem. Eng. Biotechnol.* 105 (2007) 1–30.
- [3] K.T. Klasson, M.D. Ackerson, E.C. Clausen, J.L. Gaddy, Bioconversion of synthesis gas into liquid or gaseous fuels, *Enzyme Microb. Technol.* 14 (1992) 602–608.
- [4] A.M. Henstra, J. Sipma, A. Rinzema, A.J. Stams, Microbiology of synthesis gas fermentation for biofuel production, *Curr. Opin. Biotechnol.* 18 (2007) 200–206.
- [5] S. Srikrishnan, W. Chen, N.A. Da Silva, Functional assembly and characterization of a modular xylanosome for hemicellulose hydrolysis in yeast, *Biotechnol. Bioeng.* 110 (2013) 275–285.
- [6] S.L. Tsai, M. Park, W. Chen, Size-modulated synergy of cellulase clustering for enhanced cellulose hydrolysis, *Biotechnol. J.* 8 (2013) 257–261.
- [7] X. Lu, H. Vora, C. Khosla, Overproduction of free fatty acids in *E. coli*: implications for biodiesel production, *Metab. Eng.* 10 (2008) 333–339.
- [8] Y.X. Huo, K.M. Cho, J.G. Rivera, E. Monte, C.R. Shen, Y. Yan, J.C. Liao, Conversion of proteins into biofuels by engineering nitrogen flux, *Nat. Biotechnol.* 29 (2011) 346–351.
- [9] H. Li, P.H. Oppenorth, D.G. Wernick, S. Rogers, T.Y. Wu, W. Higashide, P. Malati, Y.X. Huo, K.M. Cho, J.C. Liao, Integrated electromicrobial conversion of CO₂ to higher alcohols, *Science* 335 (2012) 1596.
- [10] B.Q. Xu, K.Q. Sun, Q.M. Zhu, W.M.H. Sachtler, Unusual selectivity of oxygenate synthesis: formation of acetic acid from syngas over unpromoted Rh in NaY zeolite, *Catal. Today* 63 (2000) 453–460.
- [11] J.F. Knifton, Syngas reactions: IX. Acetic acid from synthesis gas, *J. Catal.* 96 (1985) 439–453.
- [12] T.H. Fleisch, R.A. Sills, M.D. Briscoe, Emergence of the gas-to-liquids industry: a review of global GTL developments, *Journal of Natural Gas Chemistry* 11 (2002) 1–14.
- [13] H. Cheung, R.S. Tanke, G.P. Torrence, Acetic acid, in: *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, 2000.
- [14] P.C. Munasinghe, S.K. Khanal, Biomass-derived syngas fermentation into biofuels: opportunities and challenges, *Bioresour. Technol.* 101 (2010) 5013–5022.
- [15] R.A. Periana, O. Mironov, D. Taube, G. Bhalla, C.J. Jones, Catalytic, oxidative condensation of CH₄ to CH₃COOH in one step via CH activation, *Science* 301 (2003) 814–818.
- [16] Z. Ruan, M. Zanotti, X. Wang, C. Ducey, Y. Liu, Evaluation of lipid accumulation from lignocellulosic sugars by *Mortierella isabellina* for biodiesel production, *Bioresour. Technol.* 110 (2012) 198–205.
- [17] Y.H. Zhang, Reviving the carbohydrate economy via multi-product lignocellulose biorefineries, *J. Ind. Microbiol. Biotechnol.* 35 (2008) 367–375.
- [18] D. Mohan, C.U. Pittman, P.H. Steele, Pyrolysis of wood/biomass for bio-oil: a critical review, *Energy Fuels* 20 (2006) 848–889.
- [19] Y. Chen, J.J. Cheng, K.S. Creamer, Inhibition of anaerobic digestion process: a review, *Bioresour. Technol.* 99 (2008) 4044–4064.
- [20] P.H. Nguyen, P. Kuruparan, C. Visvanathan, Anaerobic digestion of municipal solid waste as a treatment prior to landfill, *Bioresour. Technol.* 98 (2007) 380–387.
- [21] L.T. Angenent, K. Karim, M.H. Al-Dahhan, B.A. Wrenn, R. Domiguez-Espinos, Production of bioenergy and biochemicals from industrial and agricultural wastewater, *Trends Biotechnol.* 22 (2004) 477–485.
- [22] R. Datta, Acidogenic fermentation of lignocellulose—acid yield and conversion of components, *Biotechnol. Bioeng.* 23 (1981) 2167–2170.
- [23] Z. Chi, Y. Zheng, J. Ma, S. Chen, Oleaginous yeast *Cryptococcus curvatus* culture with dark fermentation hydrogen production effluent as feedstock for microbial lipid production, *Int. J. Hydrogen Energy* 36 (2011) 9542–9550.
- [24] G. Christophe, J.L. Deo, V. Kumar, R. Nouaille, P. Fontanille, C. Larroche, Production of oils from acetic acid by the oleaginous yeast *Cryptococcus curvatus*, *Appl. Biochem. Biotechnol.* 167 (2012) 1270–1279.
- [25] J. Lian, M. Garcia-Perez, R. Coates, H. Wu, S. Chen, Yeast fermentation of carboxylic acids obtained from pyrolytic aqueous phases for lipid production, *Bioresour. Technol.* 118 (2012) 177–186.
- [26] C.K. Chen, H.P. Blaschek, Effect of acetate on molecular and physiological aspects of *Clostridium beijerinckii* NCIMB 8052 solvent production and strain degeneration, *Appl. Environ. Microbiol.* 65 (1999) 499–505.
- [27] F. Ganganella, S.U. Kuk, H. Morgan, J. Wiegel, *Clostridium thermobutyricum*: growth studies and stimulation of butyrate formation by acetate supplementation, *Microbiol. Res.* 157 (2002) 149–156.
- [28] A. Schirmer, M.A. Rude, X. Li, E. Popova, S.B. del Cardayre, Microbial biosynthesis of alkanes, *Science* 329 (2010) 559–562.
- [29] R.M. Lennen, D.J. Braden, R.A. West, J.A. Dumesic, B.F. Pfeifer, A process for microbial hydrocarbon synthesis: overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes, *Biotechnol. Bioeng.* 106 (2010) 193–202.
- [30] E.J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S.B. Del Cardayre, J.D. Keasling, Microbial production of fatty-acid-derived fuels and chemicals from plant biomass, *Nature* 463 (2010) 559–562.
- [31] A.J. Roe, C. O'Byrne, D. McLaggan, I.R. Booth, Inhibition of *Escherichia coli* growth by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity, *Microbiology* 148 (2002) 2215–2222.
- [32] G.W. Luli, W.R. Strohl, Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations, *Appl. Environ. Microbiol.* 56 (1990) 1004–1011.
- [33] J. Sambrook, D. Russell, *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York, 2001.
- [34] W. Zha, S.B. Rubin-Pitel, Z. Shao, H. Zhao, Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering, *Metab. Eng.* 11 (2009) 192–198.
- [35] P.K. Ajikumar, W.H. Xiao, K.E. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T.H. Phon, B. Pfeifer, G. Stephanopoulos, Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*, *Science* 330 (2010) 70–74.
- [36] T.A. Voelker, H.M. Davies, Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase, *J. Bacteriol.* 176 (1994) 7320–7327.
- [37] N. Aldai, K. Osoro, L.J. Barron, A.I. Najera, Gas-liquid chromatographic method for analysing complex mixtures of fatty acids including conjugated linoleic acids (cis9trans11 and trans10cis12 isomers) and long-chain (n-3 or n-6) polyunsaturated fatty acids. Application to the intramuscular fat of beef meat, *J. Chromatogr. A* 1110 (2006) 133–139.
- [38] L. You, L. Page, X. Feng, B. Beria, H.B. Pakrasi, Y.J. Tang, Metabolic pathway confirmation and discovery through (¹³C)-labeling of proteino-genic amino acids, *J. Vis. Exp.* (2012) e3583.
- [39] S.A. Wahl, M. Dauner, W. Wiechert, New tools for mass isotopomer data evaluation in (¹³C) flux analysis: mass isotope correction, data consistency checking, and precursor relationships, *Biotechnol. Bioeng.* 85 (2004) 259–268.
- [40] H. Rughoonundun, C. Granda, R. Mohee, M.T. Holtzapfel, Effect of thermochemical pretreatment on sewage sludge and its impact on carboxylic acids production, *Waste Manage.* 30 (2010) 1614–1621.
- [41] H. Rughoonundun, R. Mohee, M.T. Holtzapfel, Influence of carbon-to-nitrogen ratio on the mixed-acid fermentation of wastewater sludge and pretreated bagasse, *Bioresour. Technol.* 112 (2012) 91–97.
- [42] A.J. Wolfe, The acetate switch, *Microbiol. Mol. Biol. Rev.* 69 (2005) 12–50.
- [43] H. Lin, N.M. Castro, G.N. Bennett, K.Y. San, Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering, *Appl. Microbiol. Biotechnol.* 71 (2006) 870–874.
- [44] P. Handke, S.A. Lynch, R.T. Gill, Application and engineering of fatty acid biosynthesis in *Escherichia coli* for advanced fuels and chemicals, *Metab. Eng.* 13 (2011) 28–37.
- [45] T. Liu, C. Khosla, Genetic engineering of *Escherichia coli* for biofuel production, *Annu. Rev. Genet.* 44 (2010) 53–69.
- [46] M. Li, X. Zhang, A. Agrawal, K.Y. San, Effect of acetate formation pathway and long chain fatty acid CoA-ligase on the free fatty acid production in *E. coli* expressing acy-ACP thioesterase from *Ricinus communis*, *Metab. Eng.* 14 (2012) 380–387.
- [47] P. Xu, S. Ranganathan, Z.L. Fowler, C.D. Maranas, M.A. Koffas, Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA, *Metab. Eng.* 13 (2011) 578–587.
- [48] S. Ranganathan, T.W. Tee, A. Chowdhury, A.R. Zomorodi, J.M. Yoon, Y. Fu, J.V. Shanks, C.D. Maranas, An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*, *Metab. Eng.* 14 (2012) 687–704.

- [49] M.S. Davis, J. Solbiati, J.E. Cronan Jr., Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in *Escherichia coli*, *J. Biol. Chem.* 275 (2000) 28593–28598.
- [50] G. Stephanopoulos, A. Aristidou, J. Nielsen, *Metabolic Engineering: Principles and Methodologies*, Academic Press, 1998.
- [51] U. Sauer, F. Canonaco, S. Heri, A. Perrenoud, E. Fischer, The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*, *J. Biol. Chem.* 279 (2004) 6613–6619.
- [52] S. Atsumi, T. Hanai, J.C. Liao, Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels, *Nature* 451 (2008) 86–89.
- [53] B.B. Bond-Watts, R.J. Bellerose, M.C. Chang, Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways, *Nat. Chem. Biol.* 7 (2011) 222–227.
- [54] P.P. Peralta-Yahya, M. Ouellet, R. Chan, A. Mukhopadhyay, J.D. Keasling, T.S. Lee, Identification and microbial production of a terpene-based advanced biofuel, *Nat. Commun.* 2 (2011) 483.
- [55] C.B. Granda, M.T. Holtzapfle, Biorefineries for solvents: the MixAlco process, in: J.D. Wall, C.S. Harwood, A. Demain (Eds.), *Bioenergy*, ASM Press, Washington, DC, 2008, pp. 347–360.
- [56] Z. Liu, Z. Ruan, Y. Xiao, Y. Yi, Y.J. Tang, W. Liao, Y. Liu, Integration of sewage sludge digestion with advanced biofuel synthesis, *Bioresour. Technol.* 132 (2013) 166–170.
- [57] P. Xu, Q. Gu, W. Wang, L. Wong, A.G. Bower, C.H. Collins, M.A. Koffas, Modular optimization of multi-gene pathways for fatty acids production in *E. coli*, *Nat. Commun.* 4 (2013) 1409.
- [58] C. Dellomonaco, J.M. Clomburg, E.N. Miller, R. Gonzalez, Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals, *Nature* 476 (2011) 355–359.
- [59] F. Zhang, M. Ouellet, T.S. Batth, P.D. Adams, C.J. Petzold, A. Mukhopadhyay, J.D. Keasling, Enhancing fatty acid production by the expression of the regulatory transcription factor FadR, *Metab. Eng.* 14 (2012) 653–660.
- [60] J.T. Youngquist, R.M. Lennen, D.R. Ranatunga, W.H. Bothfeld, W.D. Marner 2nd, B.F. Pfleger, Kinetic modeling of free fatty acid production in *Escherichia coli* based on continuous cultivation of a plasmid free strain, *Biotechnol. Bioeng.* 109 (2012) 1518–1527.
- [61] X. Liu, J. Sheng, R. Curtiss 3rd, Fatty acid production in genetically modified cyanobacteria, *Proc. Natl. Acad. Sci. U.S.A.* 108 (2011) 6899–6904.

Appendix 3: Evaluating factors that influence microbial synthesis yields by regression with numerical and categorical variables.

Evaluating Factors That Influence Microbial Synthesis Yields by Linear Regression with Numerical and Ordinal Variables

Peter F. Colletti, Yogesh Goyal, Arul M. Varman, Xueyang Feng, Bing Wu, Yinjie J. Tang

Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, Missouri 63130; telephone: 314-935-3441; fax: 314-935-7211;

e-mail: yinjie.tang@seas.wustl.edu

Received 3 August 2010; revision received 18 October 2010; accepted 20 October 2010

Published online 12 November 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bit.22996

ABSTRACT: In the production of chemicals via microbial fermentation, achieving a high yield is one of the most important objectives. We developed a statistical model to analyze influential factors that determine product yield by compiling data obtained from engineered *Escherichia coli* developed within last 10 years. Using both numerical and ordinal variables (e.g., enzymatic steps, cultivation conditions, and genetic modifications) as input parameters, our model revealed that cultivation modes, nutrient supplementation, and oxygen conditions were the three significant factors for improving product yield. Generally, the model showed that product yield decreases as the number of enzymatic steps in the biosynthesis pathway increases (7–9% loss of yield per enzymatic step). Moreover, overexpression of enzymes or removal of competitive pathways (e.g., knockout) does not necessarily result in an amplification of product yield (P -value >0.1), possibly because of limited capacity in the biosynthesis pathway to accommodate an increase in flux. The model not only provides general guidelines for metabolic engineering and fermentation processes, but also allows a priori estimation and comparison of product yields under designed cultivation conditions.

Biotechnol. Bioeng. 2011;108: 893–901.

© 2010 Wiley Periodicals, Inc.

KEYWORDS: enzymatic steps; *Escherichia coli*; flux; nutrients; overexpression; P -value

Introduction

In light of rising interests and investments in green biotechnology, numerous efforts have focused on the metabolic engineering of microbial hosts to synthesize

Peter F. Colletti and Yogesh Goyal contributed equally to this work.

Yogesh Goyal's present address is Department of Chemical Engineering, Indian Institute of Technology, Gandhinagar, Gujarat 382424, India.

Correspondence to: Yinjie J. Tang

Contract grant sponsor: NSF Career Grant

Contract grant number: MCB0954016

Additional Supporting Information may be found in the online version of this article.

pharmaceuticals and biofuels (Jarboe et al., 2010; Stephanopoulos et al., 1998). For economical production of a target chemical compound, it is crucial for a microbial process to achieve high yield, titer, and productivity. Therefore, an array of engineering strategies has been routinely used for improving production, including introduction of heterologous genes, modulation of gene expression, protein modification, and high-throughput phenotype screen methods (Clomburg and Gonzalez, 2010). Optimizations of fermentation characteristics such as culture medium, pH control, temperature, and oxygen conditions are also equally important.

However, it is generally difficult to identify rate-limiting components in a complex cellular metabolism in order to improve microbial performance (Stephanopoulos et al., 2004). To this end, systems biology-based models have been developed to provide useful information for rationally engineering microbial hosts with the desired phenotype as well as to design optimal fermentation conditions (Blazek and Alper, 2010; Boghigian et al., 2010; Feist et al., 2010; Meadows et al., 2010). Cell-wide metabolic analysis via fluxomics and metabolic control theories are often used to estimate metabolic potential, product yield, nutrient limitations, and gene targets for metabolic engineering (Feist et al., 2010; Wildermuth, 2000). For example, to advance the industrial application of *Escherichia coli*, genome-scale flux balance analysis integrates a series of physical, chemical, and biological characteristics (e.g., thermodynamic directionality and energy balance) to evaluate theoretical yields for multiple native products from different feed stocks. Nevertheless, it is still difficult to reliably predict the yield produced by a culture from a given set of genetic changes and cultivation parameters. To address this, the present study proposes an empirical model based on recently published microbial production data to provide insight into the important parameters for yield optimization. We construct a linear regression model which accounts for both numerical and ordinal variables to

investigate their effects on metabolic yield. Without a priori knowledge of enzyme kinetics, biomass growth, or the stoichiometry of metabolic reactions, the model is able to estimate and compare microbial “chassis” for product synthesis. Our model provides general guidelines for pathway modifications, bioprocess optimization, and cost-estimation in industrial biotechnology.

Model Development

The construction of the empirical model focuses on *E. coli* because it is one of the most common “industrial workhorses.” It is capable of utilizing a variety of carbon sources (such as glucose, xylose, glycerol, and fatty acids) and is easily manipulated by current recombinant DNA technology. Table I lists the key influential factors that may control product yields from *E. coli*. Product biosynthesis requires a “common” route to generate metabolic intermediates such as pyruvate and acetyl-CoA via central metabolism (i.e., glycolysis, TCA cycle, and pentose phosphate pathway). A “specialized” pathway then converts the common precursors into the compounds of interest. In our model, we first divide metabolic steps into two separate numerical variables. The first variable (primary pathway) is the number of steps from the chosen carbon source to the appropriate common metabolic precursor in the central metabolism; the second variable (secondary pathway) counts the number of enzymatic steps from the common precursors to the final product of interest (Fig. 1). The choice to use two distinct variables is based on the fact that

cellular central metabolisms often show high fluxes and a robust nature for synthesizing central metabolites (e.g., pyruvate and TCA acids), while the metabolic fluxes through secondary pathways to more complex chemicals are relative small. If multiple carbon sources are used in the medium, we count the number of enzymatic steps from the carbon source which is the fewest number of steps away from the desired product.

To further develop the model, the remaining factors listed in Table I are described as ordinal-dependent variables. Each ordinal variable has either two or three categories with an intrinsic ordering to the categories. For example, nutrient conditions are divided into two categories (low and high nutrient supplements). Table II shows the categories for each ordinal-dependent variable. The relationship between model predicted yield (on a logarithmic scale) and the independent variables (regressors) can be correlated using linear regression (Tang et al., 2005), such that:

$$\begin{aligned} \log_{10} Y = & \beta_0 + \beta_{\text{PRI}} \text{PRI} + \beta_{\text{SEC}} \text{SEC} \\ & + \beta_{\text{MOD},\text{C}_2} \text{MOD}_{\text{C}_2} + \beta_{\text{MOD},\text{C}_3} \text{MOD}_{\text{C}_3} \\ & + \beta_{\text{KNO},\text{C}_2} \text{KNO}_{\text{C}_2} + \beta_{\text{NUT},\text{C}_2} \text{NUT}_{\text{C}_2} \\ & + \beta_{\text{CUL},\text{C}_2} \text{CUL}_{\text{C}_2} + \beta_{\text{OXY},\text{C}_2} \text{OXY}_{\text{C}_2} \\ & + \beta_{\text{TMP},\text{C}_2} \text{TMP}_{\text{C}_2} \end{aligned} \quad (1)$$

where $\log_{10} Y$ is the response variable. Y represents the yield (mol C in product/mol C in primary substrate), given the set of independent variables i (i.e., influential factors). β_i with

Table I. Review of influential factors affecting final product yield by *Escherichia coli*.

Factors	Mechanisms and implications	Refs.
Number of primary and secondary metabolic steps	The overall pathways for the microbial biosynthesis can be separated into two elements. The first element encompasses the pathways required to generate key metabolic intermediates from a carbon source, while the second element includes the pathways which convert these intermediates to the product	Clomburg and Gonzalez (2010)
Extent of pathway improvement	Genetic engineering, such as overexpression of native or introduction of non-native enzymes, can increase the efficiency of product biosynthesis and allow production of new compounds	Jarboe et al. (2010) and Stephanopoulos et al. (1998)
Removal of competitive pathways	Strategic deletion of specific genes or down-regulation of competitive pathways can re-direct fluxes to targeted pathways and reduce the accumulation of toxic metabolites	Bailey (1991) and Clark and Blanch (1997)
Nutrients	Cell growth requires various nutrients. Addition of extra carbon sources, intermediates or precursor compounds can improve the productivity	Pelczar et al. (1993)
Cultivation modes	Shaking flasks do not provide controlled pH and temperature conditions; a bioreactor regulates these parameters. In addition, a bioreactor provides good agitation and mixing conditions. Fed-batch operation prevents substrate inhibition which may occur if all of the nutrients are added simultaneously	Clark and Blanch (1997)
Oxygen conditions	Respiration supports high biomass growth rate but not necessarily high yield. Some products have to be fermented under low oxygen conditions. Fermentative conditions often result in the production of mixed acids and alcohol	Clark and Blanch (1997) and Pelczar et al. (1993)
Temperatures	Temperature affects enzyme activity by the Arrhenius law. An increase in temperature may also result in a decrease in dissolved oxygen content in the culture	Pelczar et al. (1993)

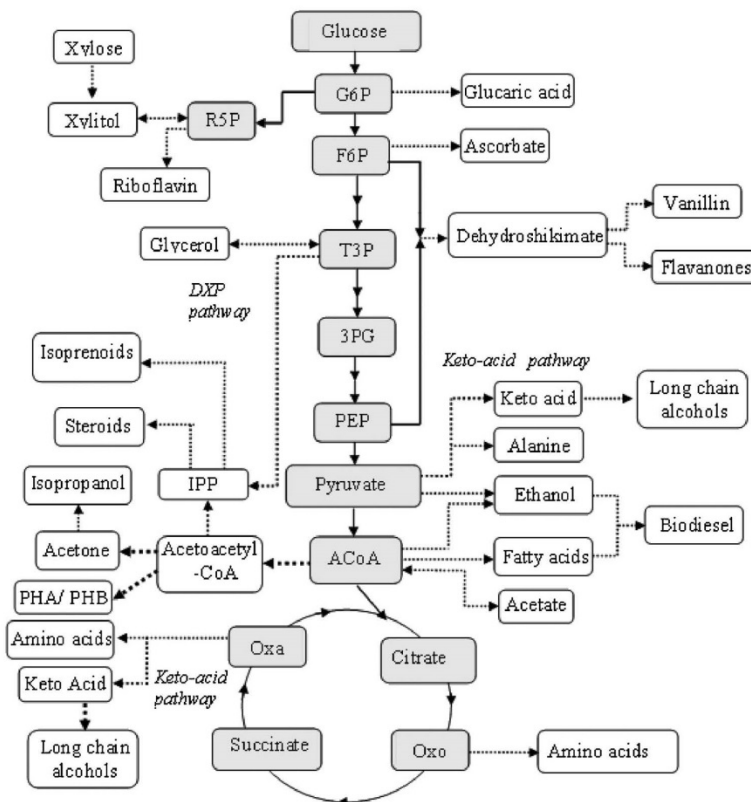


Figure 1. Simplified metabolic pathways for product synthesis. Solid arrows (—) represents primary pathways; dotted arrows represents secondary pathways (check Supplemental Fig. 1 for a more detailed pathway map). 3PG, 3-phosphoglycerate; AcCoA, acetyl-CoA; F6P, fructose 6-phosphate; G6P, glucose 6-phosphate; T3P, glyceraldehyde 3-phosphate; IPP, isopentenyl pyrophosphate; Oxa, oxaloacetate; Oxo, 2-oxoglutarate; PEP, phosphoenolpyruvate; PHA, polyhydroxy acids; PHB, poly(3-hydroxybutyrate); R5P, ribose 5-phosphate.

the exception of β_0 , is the regression coefficient of a variable i (numerical or ordinal variable), which reflects the contribution of the corresponding variable to the final product yield. The ordinal variables are assigned a value of 1 if and only if the condition fits the category in Table II; otherwise, that variable is assigned a value of 0 (Weisberg, 1985). β_0 is the intercept, which represents the combined contribution of the Category 1 of all ordinal variables. β_0 is given by Equation (2):

$$\beta_0 = \beta_{\text{MOD,C1}} + \beta_{\text{KNO,C1}} + \beta_{\text{NUT,C1}} + \beta_{\text{CUL,C1}} + \beta_{\text{OXY,C1}} + \beta_{\text{TMP,C1}} \quad (2)$$

In the final step, we compile a comprehensive database for various compounds (including isoprenoids, biofuels,

flavonones, amino acids, etc.) that have been produced using *E. coli* from 40 recently published articles (Supplementary Table I). Included in this table are the categories assigned to the experimental conditions and the reported product yield. The software package R (Team, 2009) is used to find the regression coefficients and P -values for Equation (1). A particular coefficient is considered statistically significant (90% probability) if the P -value is below 0.1.

Results

We analyze the contributions of both numerical and ordinal parameters to product yield by examining the regression model derived from the compiled data (Supplementary Table I). In general, we only select the best yields achieved by each article for the regression model (i.e., choose “the

Table II. Variables used in the linear regression model.

Notation	Category 1 (subscript C1)	Category 2 (subscript C2)	Category 3 (subscript C3)
Categorical variables			
MOD: pathway improvement (i.e., overexpression)	No genetic modification of biosynthetic pathway	Moderate pathway modification (<3 genes)	Extensive pathway modification (≥ 3 genes)
KNO: removal of competitive pathways (i.e., gene knockout)	No competitive pathway is removed	At least one competitive pathway is inhibited or removed	
NUT: nutrient sources	Fermentation occurs in defined medium (only includes trace amounts of amino acids or vitamins)	Fermentation occurs in a rich medium (nutrient carbon source >10% of primary carbon source)	
CUL: cultivation modes	Fermentation occurs in a shaking flask	Fermentation occurs in a well-controlled bioreactor	
OXY: oxygen conditions	Fermentation occurs in aerobic conditions	Fermentation occurs under oxygen-limited conditions	
TMP: fermentation temperatures	Fermentation occurs at 34°C or cooler	Fermentation occurs above 34°C	

Notation—numerical variables: steps from substrate to product.

PRI: number of reaction steps in central metabolic pathways (solid arrow in Fig. 1).

SEC: number of reaction steps in secondary metabolic pathways (dotted arrow in Fig. 1).

highest yield” under each unique production condition). The regression is based on 155 yield data (in logarithmic scale) of 36 chemicals from different carbon substrates (e.g., glucose, glycerol, etc.). From these data points, we obtain the following coefficients for Equation (1):

$$\log_{10} Y = -0.88 - 0.031 \text{PRI} - 0.041 \text{SEC} - 0.22 \text{MOD}_{C2} - 0.32 \text{MOD}_{C3} - 0.018 \text{KNO}_{C2} + 0.28 \text{NUT}_{C2} + 0.46 \text{CUL}_{C2} + 0.35 \text{OXY}_{C2} + 0.018 \text{TMP}_{C2} \quad (3)$$

Figure 2 shows a plot of the product yields obtained experimentally against those predicted by the model for the

corresponding conditions. The correlation of this model to the 155-point dataset is fair, with an R^2 value of 0.46 that represents the uncertainty attributed to the sample size and distribution. The magnitude of each coefficient indicates the extent to which the engineering of the organism and the experimental conditions affect the yield of desired product. The coefficient values must be interpreted in conjunction with the standard error and the P -value. Here we have used a P -value of 0.1 (>90% probability) as the limit below which we have considered the results to be significant (du Prel et al., 2009). Out of the nine variables used, we find that six of them have P -value of ≤ 0.1 (Table III). To ensure that the generated model is statistically sound, we have examined the

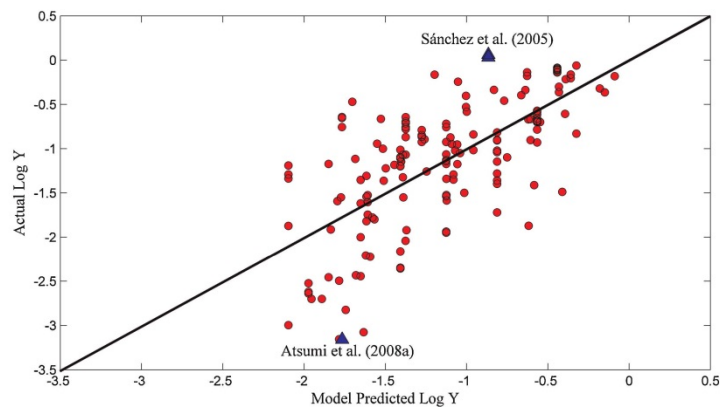


Figure 2. Distribution of actual $\log_{10} Y$ versus model generated $\log_{10} Y$ for *Escherichia coli*. The line drawn is one-to-one and passes through the origin. The R^2 value for the model is 0.46. (▲) Corresponds to the two articles discussed in the main text. [Color figure can be seen in the online version of this article, available at <http://wileyonlinelibrary.com/bit>]

Table III. Regression coefficients for the linear model comprising of eight factors (six factors are ordinal and two factors are numerical).

	Category 2			Category 3		
	Coefficient	Standard error	P-value	Coefficient	Standard error	P-value
Ordinal variables						
MOD	-0.22	0.15	0.16	-0.32	0.16	0.040
KNO	-0.018	0.11	0.88			
NUT	0.28	0.10	0.0054			
CUL	0.46	0.12	0.0036			
OXY	0.35	0.11	0.0027			
TMP	0.018	0.12	0.88			
Numerical variables						
PRI	-0.031	0.017	0.063			
SEC	-0.041	0.015	0.007			

Intercept: -0.88 ; standard error: 0.24 ; P-value: 0.00036 ; R^2 : 0.46 .

standardized residuals of the data. The standardized residual is given by the quotient of the residual of a value and its standard deviation. We found that 96% of the standardized residuals of the dataset lie between -2 and 2 . Therefore, the multiple linear regression model is statistically significant and can be a valuable tool to predict yield (Anderson et al., 2009; Montgomery et al., 2006). A plot of the standardized residual values can be found in the supplementary data (Supplementary Fig. 2).

The obtained value of β_0 (-0.88 ± 0.24) corresponds to a product yield when all of the ordinal variables for Category 2 and Category 3 are set equal to zero. The intercept β_0 reflects the yield obtained from these base conditions. To calculate product yield under actual conditions, we have to assign variables for influential factors (Supplementary Table I). For example, to predict ethanol synthesis by a wild-type strain in an anaerobic bioreactor, we can choose the following variables (Table II): PRI = 10; SEC = 2; MOD = KNO = NUT = Category 1; CUL = OXY = TMP = Category 2. The model predicts:

$$\begin{aligned} \text{Yield} &= 10^{-0.88 + (-0.031 \times 10) + (-0.041 \times 2) + 0.46 + 0.35 + 0.018} \\ &= 0.36 \end{aligned} \quad (4)$$

In this example, the result suggests that the majority of carbon supplied to the culture is used for the production of biomass and waste metabolites.

Six factors fit our criterion for significance: PRI, SEC, CUL_{C2}, OXY_{C2}, MOD_{C3}, and NUT_{C2}. As expected, the values of PRI (-0.031) and SEC (-0.041) are negative. The coefficients of these numerical variables represent the amount of decrease in log-yield for every additional enzymatic step: each enzyme may reduce the carbon yield by 7% ($= 10^{-0.031}$) in the primary pathway and by 9% ($= 10^{-0.041}$) in the secondary pathway. The two coefficients represent the loss of product yield that can accumulate as the metabolic pathway gets longer. Therefore, products from primary pathways such as pyruvate and succinate may have a yield close to the theoretical limit, but the yield of

products requiring many enzyme steps will deviate significantly from the theoretical output because of imperfect enzyme specificity and activity. We note that fermentation using a well-controlled bioreactor (reducing substrate inhibition, pH perturbation, oxygen maintenance, etc.) can improve the yield by 2.9 times ($= 10^{0.47}$) when compared to a shake-flask mode. Meanwhile, fed-batch culture allows the use of high concentrations of substrates so that the product titer can be also improved. Under anaerobic or microaerobic conditions, the yield is increased by a factor of 2.2 ($= 10^{0.35}$) compared to aerobic conditions. These low oxygen conditions, however, limit biomass growth and metabolic reaction rates.

On the other hand, the effect of temperature on the product yield is not significant because most culture temperatures are between 30 and 37°C . In this range, the difference in yield between Category 1 ($\leq 34^\circ\text{C}$) and Category 2 ($> 34^\circ\text{C}$) is minimal ($4\% = 10^{-0.018}$). Furthermore, the coefficients for MOD_{C2} and MOD_{C3} are both negative, suggesting that genetic modification of pathways may result in a decreased yield compared to the base condition where the pathway native to the host is employed. Since the P-value for MOD_{C2} is 0.16 , there is a probability that a moderate genetic modification may still result in an increased yield. The P-value for MOD_{C3} is 0.04 , which indicates that extensive pathway engineering will possibly result in a low yield ($> 95\%$) in *E. coli* even for carefully designed genetic strategies. There are a few explanations for this observation. Many studies involve the importation and expression of multiple heterologous genes for biosynthesis, and non-native pathways may be incompatible or inferior to the native pathway in the host, which may lead to lower biosynthetic efficiency. Second, pathway engineering often induces metabolic imbalance and accumulation of toxic metabolites (Atsumi et al., 2008b). Third, the initial rate-limiting enzymes in the pathway may shift to other enzymes after overexpression because of inherent low enzyme activities through the entire pathway, that is, limited capacity in the biosynthesis pathway cannot accommodate flux increase (Leonard et al., 2010).

Without a doubt, gene knockout is a useful tool to improve product yield (especially for biosynthesis of metabolites in the central metabolism, e.g., succinate) (Sánchez et al., 2006). However, we also note that the coefficient of KNO_{C_2} is slightly negative (-0.018) which means that removal of competitive pathways (i.e., genetic knockout) may not successfully improve the biosynthesis yield. While this conclusion seems highly uncertain (P -value = 0.88; standard error = 0.11), it has been shown via in silico analysis that knockout strategies cannot ensure the improvement for product yield even though it is expected to channel more carbon to the final product (Blazcek and Alper, 2010; Boghigian et al., 2010; Feist et al., 2010; Meadows et al., 2010). This inconsistency can be attributed to unfavorable metabolite accumulation and low capacity of biosynthesis pathways for flux amplification after removal of competitive pathways. For example, deletion of the acetate biosynthesis pathway did not improve the productivity of acetyl-CoA (an intermediate for many bioproducts), but rather resulted in an accumulation of pyruvate (Tomar et al., 2003). Furthermore, it has been shown that invariability of metabolic flux under mutagenic genotypes is an important feature in biological systems, and down-regulation of certain pathways may inhibit cell growth and cause imbalance in the energy metabolism. Such self-regulation of metabolism minimizes the performance of knockout-strains and makes the outcome from knockout strategies uncertain.

Discussion

The regression model presented in this article permits a priori calculation of the potential product yield from engineered *E. coli*. The model is developed through a mathematical and conceptual understanding of metabolic processes. Every variable, except for the primary and secondary number of metabolic steps, is categorized to account for the vast number of reported experimental conditions. The coefficients from the multiple linear regressions suggest several guidelines for improving final product yield, such as choosing the carbon substrate which has the shortest pathway to the final products, adding a proper nutrient source, improving cultivation modes, and utilizing pathways native to the host. Although application of these strategies can improve the conversion of carbon sources to product, reduce waste metabolites, and alleviate inhibitions due to toxic metabolites, such strategies have to be balanced with economical considerations. Furthermore, current genetic modification technologies may have unpredictable results when applied for the improvement of microbial performance. For example, knockout of competitive pathways (Bailey, 1991; Lee et al., 2005, 2007; Park et al., 2007; Trinh et al., 2006, 2008) or expression of heterologous genes (Bailey, 1991; Choi et al., 2010; Park et al., 2007; Zhang et al., 2007) are expected to shift metabolic fluxes to the desired product by removing a rate-limiting step or permitting the use of an

alternative carbon source. However, the actual behavior of a designed biosynthetic pathway can vary significantly, especially because the optimal impact of a heterologous enzyme on production improvement is not guaranteed due to its potential "side effects" such as metabolic imbalance, cytotoxicity, or incompatibility with the host organism (Atsumi et al., 2008b). Therefore, if we are to improve current metabolic engineering strategies, it is important to develop systems biology in order to understand unpredictable metabolic regulations and control behavior.

Although this empirical model provides a prediction of product yield under different conditions, the model may only be approximate. First, the model cannot determine if the final product is toxic or if the imported non-native pathway is compatible to the host. For example, the yield of 1-butanol production is low (yield = 0.0007, 20-fold below the model predicted yield) for engineered *E. coli* with a synthetic pathway from *Clostridium* (Atsumi et al., 2008a). As a comparison, production of isobutanol (less toxic than 1-butanol) via a native keto-acid pathway shows a yield higher than the model prediction (Atsumi et al., 2008b). Second, the yield calculation does not include all carbon sources such as nutrient supplements and CO_2 . For example, very high succinate yield (above 100%) from glucose was achieved because an anaplerotic pathway could fix CO_2 as a secondary carbon source (Sánchez et al., 2005) (Fig. 2). Furthermore, the category defining nutrient conditions is imprecise since most experiments involve the addition of nutrient-rich supplements (e.g., yeast extract) in varying amounts. A few studies involved the addition of intermediate compounds or precursors to improve biosynthesis fluxes. In order to reduce the potential complexity of the model, the variable NUT_{C_2} categorizes all such nutrient conditions in a generalized manner (Table II). Therefore, the model is not able to specifically account for the contribution of special nutrients or the addition of precursors to the final yield. For example, Leonard et al. (2007) engineered pathways for flavonoid production using a minimal medium including both glucose and trace phenylpropanoic acids. The regression model predicts a calculated yield (average yield = 0.008) 2–3 times lower than the experimental yield (average yield = 0.020). Third, the model cannot predict the efficiency of a heterogeneous enzyme when employed in metabolic engineering experiments. For example, Hanai et al. (2007) incorporated two different alcohol dehydrogenases in *E. coli* strains for isopropanol production. Although the two enzymes encode for the same function, the *adh* from *Clostridium beijerinckii* produced twice as much isopropanol as the gene from *Thermoanaerobacter brockii*. Finally, other factors may also influence the model accuracy, including insufficient information to precisely calculate yield, complicated metabolic behavior and pathway regression, variation of experimental conditions and analytic methods, unequal research quality and merit between different articles, and simplification of model factors using categorical numbers.

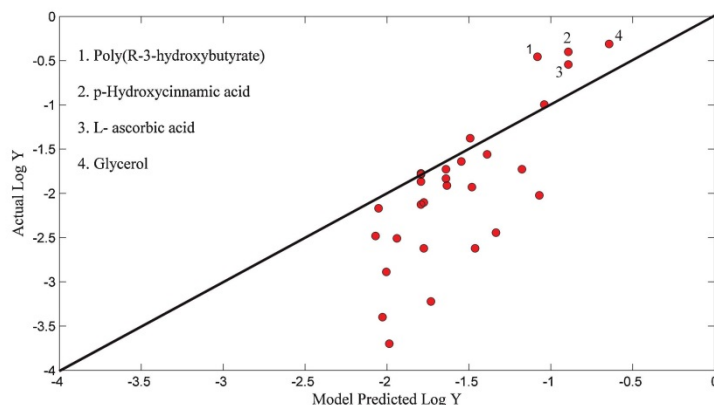


Figure 3. Distribution of actual $\log_{10} Y$ versus model generated $\log_{10} Y$ for *Saccharomyces cerevisiae*. The diagonal line drawn is one-to-one and passes through the origin. The numbered points are from the following articles: (1) (Carlson and Srienc, 2006); (2) (Vannelli et al., 2007); (3) (Sauer et al., 2004); and (4) (Overkamp et al., 2002). [Color figure can be seen in the online version of this article, available at <http://wileyonlinelibrary.com/bit>]

In an effort to examine and compare bio-production in another system, we use our model (Eq. 3, same coefficients) to analyze the yields of ~ 20 products achieved by *Saccharomyces cerevisiae* based on several recent articles. A table of the collected data for *S. cerevisiae* is available in the supplementary data (Supplementary Table II). Figure 3 shows that the product yields for *S. cerevisiae* are mostly below the predicted yield calculated from the *E. coli* model, especially for the biosynthesis of “low-yield compounds.” Notably, several data points in Figure 3 are greater than those of model predictions due to unique biosynthesis methods. First, the production of *p*-hydroxycinnamic acid from glucose (20 g/L) was unusually high (Vannelli et al., 2007). This very high yield was achieved by supplementing 10 g/L phenylalanine to enhance product synthesis. Second, a study on ascorbic acid production from galactose incorporated a two-stage process: using glucose for biomass growth, then adding galactose for product synthesis in the stationary phase (Sauer et al., 2004). Such a cultivation process could improve the yield of ascorbic acid because small carbon fluxes are directed to biomass synthesis during the product production phase. Third, a study on poly[(R)-3-hydroxybutyrate] (PHB) synthesis supplemented the culture with pantothenate (a precursor of final product) to enhance PHB production (Carlson and Srienc, 2006). This indicates that addition of precursor or intermediate compounds can be significantly more efficient for promoting yield than addition of undefined nutrient sources (e.g., yeast extract) because it only requires a small number of enzymatic steps to convert the precursor or intermediate to the final product. Fourth, a study on glycerol production using glucose produced a higher yield than the model prediction (Overkamp et al., 2002). This can be attributed to the inclusion of ethanol along with glucose in chemostat cultivation.

Furthermore, we also compared chemical production from different microbial hosts (e.g., *S. cerevisiae*, *Schizosaccharomyces pombe*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Klebsiella pneumoniae*, *Clostridium diolis*, and *Clostridium acetobutylicum*) (Fig. 4). Without requiring extensive genetic modifications, some microorganisms can synthesize certain metabolic products using their native pathways (Alper and Stephanopoulos, 2009). Data points from these environmental microbes are distributed about the one-to-one line in the figure, showing their biosynthesis capability is comparable to that of *E. coli*. The model also demonstrates that other species may possess a significant disadvantage in the production of a particular product. For example, butanol synthesis by *S. cerevisiae* is 1–2 orders of magnitude lower than the model predictions because yeast shows an inability to utilize the engineered heterogeneous genes efficiently and a low tolerance for butanol (Fig. 4), while *C. acetobutylicum* has a very robust butanol production comparable to engineered *E. coli* strains (Atsumi et al., 2008a; Sillers et al., 2008; Steen et al., 2008).

Finally, the model can evaluate new methods that are being developed to increase final product yields. An example is protein scaffolding, which can avoid channeling intermediates away from the desired product and significantly improve the product yields (Dueber et al., 2009). By plotting the actual yield and the predicted yield (Supplementary Table III), the sample point (glucaric acid) lies above the one-to-one line and the sample point (mevalonate) lies close to the one-to-one line after scaffolding (Fig. 4). This implies that protein scaffolding can be useful for maximizing biosynthetic potential in *E. coli*. Another example of a new method is the application of both metabolic and protein engineering for levopimaradiene synthesis (Leonard et al., 2010). The experimental yield (~ 0.03) is fairly close to the

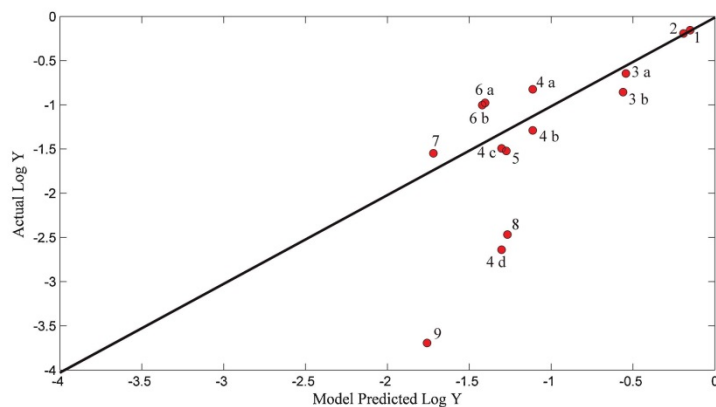


Figure 4. Distribution of actual $\log_{10} Y$ versus model generated $\log_{10} Y$ for various non-model species as well as protein scaffolding in *Escherichia coli*. The line drawn is one-to-one and passes through the origin. The numbers represents: (1) *Klebsiella pneumoniae*; 1,3-propanediol (Zhang et al., 2006), (2) *Clostridium diolis*; 1,3-propanediol (Otte et al., 2009), (3) *Clostridium acetobutylicum*; butanol without knockout (3a); butanol with knockout (3b) (Sillers et al., 2008), (4) *Escherichia coli*; glutaric acid with scaffold (4a); glutaric acid without scaffold (4b); mevalonate with scaffold (4c); mevalonate without scaffold (4d) (Dueber et al., 2009), (5) *Escherichia coli*; levopimaradiene (Leonard et al., 2010), (6) *Corynebacterium glutamicum*; isobutanol with knockout (6a); isobutanol without knockout (6b) (Smith et al., 2010), (7) *Bacillus subtilis*; riboflavin (Li et al., 2006), (8) *Schizosaccharomyces pombe*; vanillin (Hansen et al., 2009), and (9) *Saccharomyces cerevisiae*; butanol (Steen et al., 2008). [Color figure can be seen in the online version of this article, available at <http://wileyonlinelibrary.com/bit>]

predicted yield (~ 0.05), which indicates the effectiveness of their pathway engineering strategies.

Conclusions

The statistic-based model permits a priori estimation of final product yield of a metabolically produced compound based upon the genetic design of the microbial host system and the conditions under which the fermentation process occurs. This empirical model reflects the capability of current biotechnology to engineer *E. coli* to achieve the maximal biosynthetic yield. Such a model can provide information complementary to the conventional flux balance model which reflects the productivity of the “optimal” cellular metabolism. The model equation can be used to economically optimize the conditions for the production of a specific compound, providing its users with an estimate of the yield without performing actual experiments. The model can also be applied to other microbial chassis and used to compare their biosynthesis productivity to *E. coli*. The predictive power of our model, however, is limited by variations in experimental conditions and uncertainties in biological behaviors. For example, the model cannot evaluate the effect of precursor metabolites or toxic intermediates on final product yields. In general, this statistically based model provides a new method towards developing a comprehensive and quantitative understanding of experimental approaches for metabolic engineering purposes. The model can be widely useful to plan wet lab

experiments across a wide range of applications in both scientific and industrial communities.

This study was supported by an NSF Career Grant (MCB0954016) and an international undergraduate research program at the Indian Institute of Technology (directed by Dr. Sudhir K. Jain). The authors are grateful to Jean Zhang in the Division of Biostatistics at Washington University Medical School for her beneficial discussions.

References

- Alper H, Stephanopoulos G. 2009. Engineering for biofuels: Exploiting innate microbial capacity or importing biosynthetic potential? *Nat Rev Microbiol* 7(10):715–723.
- Anderson DR, Sweeney DJ, Williams TA. 2009. Essentials of statistics for business and economics, 5th edn. Mason, OH: South-Western, Cengage Learning.
- Atsumi S, Cann AF, Connor MR, Shen CR, Smith KM, Brynildsen MP, Chou KJ, Hanai T, Liao JC. 2008a. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab Eng* 10(6):305–311.
- Atsumi S, Hanai T, Liao JC. 2008b. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451(7174):86–89.
- Bailey JE. 1991. Toward a science of metabolic engineering. *Science* 252(5013):1668–1675.
- Blazeck J, Alper H. 2010. Systems metabolic engineering: Genome-scale models and beyond. *Biotechnol J* 5(7):647–659.
- Boghigian BA, Seth G, Kiss R, Pfeifer BA. 2010. Metabolic flux analysis and pharmaceutical production. *Metab Eng* 12(2):81–95.
- Carlson R, Srien F. 2006. Effects of recombinant precursor pathway variations on poly[(R)-3-hydroxybutyrate] synthesis in *Saccharomyces cerevisiae*. *J Biotechnol* 124(3):561–573.
- Choi HS, Lee SY, Kim TY, Woo HM. 2010. *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl Environ Microbiol* 76(10):3097–3105.

- Clark DS, Blanch HW. 1997. Biochemical engineering. New York: Marcel Dekker, Inc.
- Cloomb JM, Gonzalez R. 2010. Biofuel production in *Escherichia coli*: The role of metabolic engineering and synthetic biology. *Appl Microbiol Biotechnol* 86(2):419–434.
- du Prel JB, Hommel G, Rohrig B, Blettner M. 2009. Confidence interval or p-value? Part 4 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 106(19):335–339.
- Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KL, Keasling JD. 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat Biotechnol* 27(8):753–759.
- Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, Palsson BO. 2010. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metab Eng* 12(3):173–186.
- Hanai T, Atsumi S, Liao JC. 2007. Engineered synthetic pathway for isopropanol production in *Escherichia coli*. *Appl Environ Microbiol* 73(24):7814–7818.
- Hansen EH, Moller BL, Kock GR, Bunner CM, Kristensen C, Jensen OR, Okkels FT, Olsen CE, Motawia MS, Hansen J. 2009. De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Appl Environ Microbiol* 75(9):2765–2774.
- Jarboe LR, Zhang X, Wang X, Moore JC, Shanmugam KT, Ingram LO. 2010. Metabolic engineering for production of biorenewable fuels and chemicals: Contributions of synthetic biology. *J Biomed Biotechnol* 2010: 761042.
- Lee SJ, Lee DY, Kim TY, Kim BH, Lee J, Lee SY. 2005. Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and *in silico* gene knockout simulation. *Appl Environ Microbiol* 71(12):7880–7887.
- Lee KH, Park JH, Kim TY, Kim HU, Lee SY. 2007. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol Syst Biol* 3:149.
- Leonard E, Lim KH, Saw PN, Koffas MA. 2007. Engineering central metabolic pathways for high-level flavonoid production in *Escherichia coli*. *Appl Environ Microbiol* 73(12):3877–3886.
- Leonard E, Ajikumar PK, Thayer K, Xiao WH, Mo JD, Tidor B, Stephanopoulos G, Prather KL. 2010. Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc Natl Acad Sci USA* 107(31):13654–13659.
- Li XJ, Chen T, Chen X, Zhao XM. 2006. Redirection electron flow to high coupling efficiency of terminal oxidase to enhance riboflavin biosynthesis. *Appl Microbiol Biotechnol* 73(2):374–383.
- Meadows AL, Karnik R, Lam H, Forestell S, Snedecor B. 2010. Application of dynamic flux balance analysis to an industrial *Escherichia coli* fermentation. *Metab Eng* 12(2):150–160.
- Montgomery DC, Peck EA, Vining GG. 2006. Introduction to linear regression analysis. New York: John Wiley & Sons.
- Otte B, Grunwaldt E, Mahmoud O, Jennewein S. 2009. Genome shuffling in *Clostridium diolis* DSM 15410 for improved 1,3-propanediol production. *Appl Environ Microbiol* 75(24):7610–7616.
- Overkamp KM, Bakker BM, Kotter P, Luttik MA, Van Dijken JP, Pronk JT. 2002. Metabolic engineering of glycerol production in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 68(6):2814–2821.
- Park JH, Lee KH, Kim TY, Lee SY. 2007. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc Natl Acad Sci USA* 104(19):7797–7802.
- Pelczar MJ, Jr., Chan ECS, Krieg NR. 1993. Microbiology: Concepts and applications. New York: McGraw-Hall Inc.
- Sánchez AM, Bennett GN, San KY. 2005. Novel pathway engineering design of the anaerobic central metabolic pathway in *Escherichia coli* to increase succinate yield and productivity. *Metab Eng* 7(3):229–239.
- Sánchez AM, Bennett GN, San KY. 2006. Batch culture characterization and metabolic flux analysis of succinate-producing *Escherichia coli* strains. *Metab Eng* 8(3):209–226.
- Sauer M, Branduardi P, Valli M, Porro D. 2004. Production of L-ascorbic acid by metabolically engineered *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii*. *Appl Environ Microbiol* 70(10):6086–6091.
- Sillers R, Chow A, Tracy B, Papoutsakis ET. 2008. Metabolic engineering of the non-sporulating, non-solventogenic *Clostridium acetobutylicum* strain M5 to produce butanol without acetone demonstrate the robustness of the acid-formation pathways and the importance of the electron balance. *Metab Eng* 10(6):321–332.
- Smith KM, Cho KM, Liao JC. 2010. Engineering *Corynebacterium glutamicum* for isobutanol production. *Appl Microbiol Biotechnol* 87(3): 1045–1055.
- Steen EJ, Chan R, Prasad N, Myers S, Petzold CJ, Redding A, Ouellet M, Keasling JD. 2008. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microb Cell Fact* 7:36.
- Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic engineering: Principles and methodologies. San Diego: Academic Press.
- Stephanopoulos G, Alper H, Moxley J. 2004. Exploiting biological complexity for strain improvement through systems biology. *Nat Biotechnol* 22(10):1261–1267.
- Tang YJ, Qi L, Krieger-Brockett B. 2005. Evaluating factors that influence microbial phenanthrene biodegradation rates by regression with categorical variables. *Chemosphere* 59(5):729–741.
- Team RDC. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: ISBN 3-900051-07-0, URL: <http://www.R-project.org>.
- Tomar A, Eiteman MA, Altman E. 2003. The effect of acetate pathway mutations on the production of pyruvate in *Escherichia coli*. *Appl Microbiol Biotechnol* 62(1):76–82.
- Trinh CT, Carlson R, Wlaschin A, Srien F. 2006. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metab Eng* 8(6):628–638.
- Trinh CT, Unrean P, Srien F. 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Environ Microbiol* 74(12):3634–3643.
- Vannelli T, Wei Qi W, Sweigard J, Gatenby AA, Sariaslani FS. 2007. Production of p-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metab Eng* 9(2):142–151.
- Weisberg S. 1985. Applied linear regression. New York: John Wiley & Sons.
- Wildermuth MC. 2000. Metabolic control analysis: Biological applications and insights. *Genome Biol* 1(6):reviews1031.
- Zhang Y, Li Y, Du C, Liu M, Cao Z. 2006. Inactivation of aldehyde dehydrogenase: A key factor for engineering 1,3-propanediol production by *Klebsiella pneumoniae*. *Metab Eng* 8(6):578–586.
- Zhang X, Jantama K, Moore JC, Shanmugam KT, Ingram LO. 2007. Production of L-alanine by metabolically engineered *Escherichia coli*. *Appl Microbiol Biotechnol* 77(2):355–366.

Curriculum Vitae

Arul Mozhy Varman

Department of Energy, Environmental and Chemical Engineering
Washington University in Saint Louis • Campus Box 1180 • One Brookings Drive
St. Louis, MO 63130 • Email: avarman@go.wustl.edu

EDUCATION

Ph.D. in Energy, Environmental and Chemical Engineering,

Washington University in St. Louis, 2013

M.S. in Energy, Environmental and Chemical Engineering,

Washington University in St. Louis, 2010

M.Tech. in Chemical Engineering,

Indian Institute of Technology Bombay, India, 2006

B.Tech. in Chemical Engineering,

Pondicherry University, India, 2003

RESEARCH EXPERIENCE

PhD dissertation: “Systems metabolic engineering of microbial cell factories for the synthesis of value-added chemicals”, Advisor: Dr. Yinjie J. Tang, Washington University in St. Louis

- Created an engineered strain of *Synechocystis* 6803 for isobutanol synthesis.
- Conceptualized and successfully engineered the D-lactate pathway in *Synechocystis* 6803.
- Conducted fermentations and developed kinetic models for production of fatty acids and isobutanol in *E. coli*.
- Developed statistical models to evaluate factors influencing yield of biochemicals in engineered microbes.
- Performed ¹³C-Metabolic flux analysis (¹³C-MFA) for the isobutanol producing *E. coli*.

Master’s thesis: “An improved plasmid vector system for genetic engineering of *Synechocystis* sp. PCC 6803, a model cyanobacterium”, Advisor: Dr. Himadri B. Pakrasi, Washington University in St. Louis

- Developed my molecular biology experimental skills for the engineering of cyanobacteria.
- Analyzed the proteomics data for *Synechocystis* 6803 and mapped the results to KEGG pathways.

Master’s thesis: “Modeling the chemical equilibria for the extraction of actinides”, Advisor: Dr. Vinay A. Juvekar, Indian Institute of Technology Bombay

- Developed a liquid-liquid extraction system for the separation of actinides from radioactive wastes.
- Developed thermodynamic models to describe the chemical equilibria of actinide extraction.

AWARDS AND HONOURS

- Departmental **Graduate Student Research Award**, Washington University in St. Louis, 2012.
- **MAGEEP Travel Grant**, 4th International Symposium on Energy and Environment, India, 2012.
- **Student Travel Grant**, the 111th General Meeting of the American Society for Microbiology, 2011.
- Recipient of **James M. McKelvey fellowship** at Washington University in St. Louis, 2008-2009.

- Recipient of **Department of Atomic Energy (DAE) Fellowship**, 2004 – 2006.
- Selected for **Ministry of Human Resource and Development (India) fellowship**, 2004-2006.
- Secured **40th** rank with **99.36** percentile in **GATE-2004** (All India Entrance exam).
- Secured **8th** rank in **All India Combined Entrance Examination** conducted by DBT, 2003.

PUBLICATIONS

1. **Varman AM**, He L, You L, et.al. The use of ¹³C-based analysis to elucidate the intrinsic biosynthesis yields. “*Under review*”.
2. **Varman AM**, Yu Y, You L, Tang YJ. Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microbial Cell Factories*. 2013. 12:117.
3. **Varman AM**, Xiao Y, Pakrasi H, Tang YJ. Metabolic engineering of *Synechocystis* sp. strain PCC 6803 for isobutanol production. *Applied Environmental and Microbiology*. 2013. 79(3): 908-914.
4. Xiao Y, Ruan Z, Liu Z, Wu SG, **Varman AM**, Liu Y, Tang YJ. Engineering *Escherichia coli* to convert acetic acid to free fatty acids, *Biochemical Engineering Journal*. 2013. 76: 60-69.
5. **Varman AM**, Yi X, Feng X, et.al. Kinetic modeling and isotopic investigation of isobutanol fermentation by two engineered *Escherichia coli* strains. *Industrial & Engineering Chemistry Research*. 2012. 51 (49):15855–15863.
6. **Varman AM**, Xiao Y, Leonard E, Tang YJ. Statistics-based model for prediction of chemical biosynthesis yield from *Saccharomyces cerevisiae*. *Microbial Cell Factories*. 2011. 10:45.
7. Colletti P, Goyal Y, **Varman AM**, et.al. Evaluating factors that influence microbial synthesis yields by regression with numerical and categorical variables. *Biotechnology and Bioengineering*. 2011. 108(4):893-901.

BOOK CHAPTER

- **Varman AM**, He L, Tang YJ. “Chapter 13: Microbial metabolisms and cell culture models for biofuel production.” *Bioenergy: Principles and Applications*, **Invited book chapter**.

CONFERENCE PRESENTATIONS

- Hollinshead W, Hembree Z, **Varman AM**, et.al. “Utilization of anaerobically digested sludge for cultivation of an engineered cyanobacterium.” Annual Mid-American Environmental Engineering Conference, 2013, MO.
- **Varman AM**, Xiao Y, Yi Y, Pakrasi HB, Tang YJ. “Metabolic engineering of *Synechocystis* 6803 for the production of chemicals and fuels.” 11th Workshop on Cyanobacteria, 2013, MO.
- **Varman AM**, Xiao Y, Pakrasi HB, Tang YJ. “Metabolic engineering of the cyanobacterium *Synechocystis* 6803 for isobutanol biosynthesis.” Indo - US Workshop on Cyanobacteria, 2012, Lonavala, India.
- **Varman AM**, Chen A, Wu SG, Xiao Y, Yi Y, Le Y, Kang C, Tang YJ. “Advanced biofuel production from organic wastes.” 4th International Symposium on Energy and Environment, 2012, Mumbai, India.
- **Varman AM**, Xiao Y, Pakrasi HB, Tang YJ. “Metabolic engineering of the cyanobacterium *Synechocystis* 6803 for isobutanol biosynthesis.” 4th International Symposium on Energy and Environment, 2012, India.
- **Varman AM**, Xiao Y, Pakrasi HB, Tang YJ. “Metabolic engineering of the cyanobacterium *Synechocystis* 6803 for isobutanol biosynthesis.” Annual Mid-American Environmental Engineering Conference, 2012, IL.
- **Varman AM**, Xiao Y, Pakrasi HB, Tang YJ. “Metabolic engineering of the cyanobacterium *Synechocystis* 6803 for isobutanol biosynthesis.” Plant Biology Annual Retreat, 2012, October 19, Eureka, MO.

- **Varman AM**, Xiao Y, Nguyen A, Pakrasi HB, Tang YJ. “Metabolic engineering of the cyanobacterium *Synechocystis* 6803 for isobutanol biosynthesis.” The 112th ASM General Meeting, 2012, San Francisco, CA.
- Xiao Y, Feng X, Yu H, **Varman AM** and Tang YJ. “Kinetic modeling of isobutanol fermentation by recombinant *Escherichia Coli*.” The 2011 AIChE Annual Meeting, 2011, October 16-21, Minneapolis, MN
- **Varman AM**, Xiao Y, Feng X, et.al. “Statistics-based model for prediction of chemical biosynthesis yield from *Saccharomyces cerevisiae* and *E. coli*.” The 111th ASM General Meeting, 2011, LA.

TEACHING EXPERIENCE

- Teaching assistant, ChE 351 - Engineering Analysis of Chemical Systems, WUSTL.
- Teaching assistant, ChE 367 - Transport Phenomena I, WUSTL.
- Teaching assistant, EECE 595 - Energy and Buildings, WUSTL.
- Teaching assistant, CL 253 - Chemical Engineering Thermodynamics, IIT Bombay.
- Teaching assistant, CL 610 - Experimental methods, IIT Bombay.
- Mentored graduate and undergraduate students at Washington University and IIT Bombay.

AFFILIATIONS

- American Chemical Society.
- American Institute of Chemical Engineers.

INDUSTRIAL EXPERIENCE

Scientific Officer C, Bhabha Atomic Research Centre, Mumbai, India (June 2006 – July 2008).

- Performed research as well as designed extraction equipment, air lift pumps and tanks in utility section.
- Analyzed and prepared the safety report for an ongoing plant erection and commissioning project.
- Supervised the on-site plant erection and commissioning.

Process Engineer, Shasun chemicals and drugs, Pondicherry, India (May – July 2003)

- Provided optimization and trouble-shooting support for plant operations.
- Supervised, monitored and reported plant performance on a regular basis.

ADDITIONAL SKILLS

Software Packages: MATLAB, R, GAMS, HYSYS, Foxpro, HENS, Mathematica, AutoCAD.

Programming Languages: C, C++, Pascal, FORTRAN, Perl.

Spoken/Written: English, Tamil, Telugu, Hindi.

Others: Senior in Typewriting.

ACTIVITIES

- Initiated and organized students’ monthly meetings for the metabolic engineering cluster at WUSTL.
- Lead a team of students to prepare a departmental video at WUSTL.
- Performed in a skit at Washington University in St. Louis.
- Active member of World Wildlife Fund (WWF).